

# Quantum limits in optical interferometry

R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński

*Faculty of Physics,  
University of Warsaw,  
ul. Hoża 69, PL-00-681 Warszawa,  
Poland*

Non-classical states of light find applications in enhancing the performance of optical interferometric experiments, with notable example of gravitational wave-detectors. Still, the presence of decoherence hinders significantly the performance of quantum-enhanced protocols. In this review, we summarize the developments of quantum metrology with particular focus on optical interferometry and derive fundamental bounds on achievable quantum-enhanced precision in optical interferometry taking into account the most relevant decoherence processes including: phase diffusion, losses and imperfect interferometric visibility. We introduce all the necessary tools of quantum optics as well as quantum estimation theory required to derive the bounds. We also discuss the practical attainability of the bounds derived and stress in particular that the techniques of quantum-enhanced interferometry which are being implemented in modern gravitational wave detectors are close to the optimal ones.

## CONTENTS

I. Introduction	1	1. Phase diffusion	27
II. Quantum states of light	4	2. Photonic losses	27
A. Mode description	4	3. Imperfect visibility	28
B. Gaussian states	4	B. Bounds in the QFI approach	28
1. Coherent states	5	1. Imperfect visibility	29
2. Single-mode squeezed states	5	2. Photonic losses	30
3. Two-mode squeezed states	6	3. Phase diffusion	32
4. General two-mode Gaussian states	6	C. Bayesian approach	33
C. Definite photon number states	7	1. Photonic losses	33
D. Particle description	7	2. Phase diffusion	34
E. Mode vs particle entanglement	8	D. Practical schemes saturating the bounds	34
III. Mach-Zehnder interferometry	8	1. Bounds for indefinite photon number states	34
A. Phase-sensing uncertainty	9	2. Coherent + squeezed vacuum strategy	35
B. Coherent-state interferometry	10	VII. Conclusions	36
C. Fock state interferometry	10	References	37
D. Coherent + squeezed-vacuum interferometry	10		
E. Definite photon-number state interferometry	11		
F. Other interferometers	12		
IV. Estimation Theory	13		
A. Classical parameter estimation	14		
1. Fisher Information approach	14		
2. Bayesian approach	15		
3. Example: Transmission coefficient estimation	16		
B. Quantum parameter estimation	17		
1. Quantum Fisher Information approach	17		
2. Bayesian approach	19		
V. Quantum limits in decoherence-free interferometry	19		
A. Quantum Fisher Information approach	20		
B. Bayesian approach	21		
C. Indefinite photon-number states	22		
1. Role of the reference beam	22		
2. Optimal indefinite photon number strategies	23		
3. Gaussian states	23		
D. Role of entanglement	24		
E. Multi-pass protocols	25		
VI. Quantum limits in realistic interferometry	25		
A. Decoherence models	26		

## I. INTRODUCTION

Without much exaggeration one may say that optics is basically the science of light interference. Light interference effects were behind the final acceptance of classical wave optics and abandoning the Newtonian corpuscular theory of light. Conceptual insight into the process of interfering light waves prompted development of a number of measurement techniques involving well controlled interference effects and gave birth to the field of optical interferometry (Hariharan, 2003). At the fundamental level, classical interferometry is all about observing light intensity variations (intensity fringes) resulting from a change in relative phases between two (or more) overlapping light waves, e.g. when a single light beam is split into a number of paths with tunable optical path-length differences and made to interfere on the screen. The number of applications is stunning and ranges from basic length measurements via spectroscopic interferometric techniques to the most spectacular examples involv-

ing stellar interferometry and gravitational-wave detectors (Pitkin *et al.*, 2011).

Coherent properties of light as well as the degree of overlap between the interfering beams determine the visibility of the observed intensity fringes and are crucial for the quality of any interferometric measurement. Still, when asking for fundamental limitations on precision of estimating e.g. a phase difference between the arms of a Mach-Zehnder interferometer, there is no particular answer within a purely classical theory, where both the light itself as well as the detection process are treated classically. In classical theory intensity of light can in principle be measured with arbitrary precision and as such allows to detect in principle arbitrary small phase shifts in an interferometric experiment.

This, however, is no longer true when semi-classical theory is considered in which the light is still treated classically but the detection process is quantized, so that instead of a continuous intensity parameter the number of energy quanta (photons) absorbed is being measured. The absorption process within the semi-classical theory has a stochastic character and the number of photons detected obeys Poissonian or super-Poissonian statistics (Fox, 2006). If light intensity fluctuations can be neglected, the number of photons detected,  $N$ , follows the Poissonian statistics with photon number standard deviation  $\Delta N = \sqrt{\langle N \rangle}$ , where  $\langle N \rangle$  denotes the mean number of photons detected. This implies that the determination of the relative phase difference  $\varphi$  between the arms of the interferometer, based on the number of photons detected at the output ports, will be affected by the relative uncertainty  $\Delta\varphi \propto \Delta N / \langle N \rangle = 1/\sqrt{\langle N \rangle}$  referred to as the *shot noise*. The shot noise plays a fundamental role in the semi-classical theory and in many cases it is indeed the factor limiting achievable interferometric sensitivities. In modern gravitational wave detectors, in particular, the shot noise is the dominant noise term in the noise spectral density for frequencies above few hundred Hz (LIGO Collaboration, 2011, 2013; Pitkin *et al.*, 2011).

Yet, the shot noise should not be regarded as a fundamental bound whenever non-classical states of light are considered—we review the essential aspects of non-classical light relevant from the point of view of interferometry in Sec. II. The *sub-Poissonian* statistics characteristic for the so-called squeezed states of light may offer a precision enhancement in interferometric scenarios by reducing the photon number fluctuations at the output ports. First proposals demonstrated that sending coherent light together with the squeezed vacuum state into the two separate input ports of a Mach-Zehnder interferometer offers estimation precision beating the shot noise and attaining the  $1/\langle N \rangle^{2/3}$  scaling of the phase estimation precision (Caves, 1981), while consideration of more general two-mode squeezed states showed that even  $1/\langle N \rangle$  scaling is possible (Bondurant and Shapiro,

1984; Yurke *et al.*, 1986). A number of papers followed, studying in more detail the phase-measurement probability distribution and proposing various strategies leading to  $1/\langle N \rangle$  precision scaling (Braunstein, 1992; Dowling, 1998; Holland and Burnett, 1993; Sanders and Milburn, 1995). Similar observations have been made in the context of precision spectroscopy, where spin-squeezed states have been shown to offer a  $1/N$  scaling of atomic transition frequency estimation precision, where  $N$  denotes the number of atoms employed (Wineland *et al.*, 1994, 1992). Sec. III provides a detailed framework for deriving the above results. However, already at this point a basic intuition should be conveyed that one *cannot* go beyond the shot noise limit whenever an interferometric experiment may be regarded in a spirit that each photon interferes only with itself. In fact, the only possibility of surpassing this bound is to use light sources that exhibit correlations in between the constituent photons, e.g. squeezed light, so that the interference process may benefit from the properties of inter-photonic entanglement.

These early results provided a great physical insight into the possibilities of quantum enhanced interferometry and the class of states that might be of practical interest for this purpose. The papers lacked generality, however, by considering specific measurement-estimation strategies, in which the error in the estimated phase was related via a simple error-propagation formula to the variance of some experimentally accessible observable, e.g. photon number difference at the two output ports of the interferometer, or by studying the width of the peaks in the shape of the phase-measurement probability distribution.

Given a particular state of light fed into the interferometer, it is a priori not clear what is the best measurement and estimation strategy yielding the optimal estimation precision. Luckily, the tools designed to answer these kinds of questions had already been present in the literature under the name of *quantum estimation theory* (Helstrom, 1976; Holevo, 1982). The *Quantum Fisher Information* (QFI) as well as the *cost of Bayesian inference* provide a systematic way to quantify the ultimate limits on performance of phase-estimation strategies for a given quantum state, which are already optimized over all theoretically admissible quantum measurements and estimators. The concept of the QFI and the Bayesian approach to quantum estimation are reviewed in Sec. IV. As a side remark, we should note, that by treating the phase as an evolution parameter to be estimated and separating explicitly the measurement operators from the estimator function, quantum estimation theory circumvents some of the mathematical difficulties that arise if one insists on the standard approach to quantum measurements and attempts to define the *quantum phase operator* representing the phase observable being measured (Barnett and Pegg, 1992; Lynch, 1995; Noh *et al.*, 1992; Summy and Pegg, 1990).

The growth of popularity of the QFI in the field of quantum metrology was triggered by the seminal paper of Braunstein and Caves (1994) advocating the use of QFI as a natural measure of distance in the space of quantum states. The QFI allows to pin down the optimal probe states that are the most sensitive to small variations of the estimated parameter by establishing the fundamental bound on the corresponding parameter sensitivity valid for arbitrary measurements and estimators. Following these lines of reasoning the  $1/N$  bound, referred to as the *Heisenberg limit*, on the phase estimation precision using  $N$ -photon states has been claimed fundamental and the NOON states were formally proven to saturate it (Bollinger *et al.*, 1996). Due to close mathematical analogies between optical and atomic interferometry (Bollinger *et al.*, 1996; Lee *et al.*, 2002) similar bounds hold for the problem of atomic transition-frequency estimation and more generally for any arbitrary unitary parameter estimation problem, i.e. the one in which an  $N$ -particle state evolves under a unitary  $U_\varphi^{\otimes N}$ ,  $U_\varphi = \exp(-i\hat{H}\varphi)$ , with  $\hat{H}$  being a general single-particle evolution generator and  $\varphi$  the parameter to be estimated (Giovannetti *et al.*, 2004, 2006).

A complementary framework allowing to determine the fundamental bounds in interferometry is the Bayesian approach, in which one assumes the estimated parameter to be a random variable itself and explicitly defines its *prior* distribution to account for the initial knowledge about  $\varphi$  before performing the estimation. In the case of interferometry the typical choice is the flat prior  $p(\varphi) = 1/2\pi$  which reflects the initial ignorance of the phase. The search for the optimal estimation strategies within the Bayesian approach is possible thanks to the general theorem on the optimality of the *covariant measurements* in estimation problems satisfying certain group symmetry (Holevo, 1982). In the case of interferometry, a flat prior guarantees the phase-shift,  $U(1)$ , symmetry and as a result the optimal measurement operators can be given explicitly and they coincide with the eigenstates of the Pegg-Barnett “phase operator” (Barnett and Pegg, 1992). This makes it possible to optimize the strategy over the input states and for simple cost functions allows to find the optimal probe states (Berry and Wiseman, 2000; Bužek *et al.*, 1999; Luis and Perina, 1996). In particular, for the  $4\sin^2(\delta\varphi/2)$  cost function which approximates the variance for small phase deviations  $\delta\varphi$ , the corresponding minimal estimation uncertainty has been found to read  $\Delta\varphi \approx \pi/N$  for large  $N$  providing again a proof of the possibility of achieving the Heisenberg scaling, yet with an additional  $\pi$  coefficient. It should be noted that the optimal states in the above approach that have been found independently in (Berry and Wiseman, 2000; Luis and Perina, 1996; Summy and Pegg, 1990) have completely different structure to the NOON states which are optimal when QFI is considered as the figure of merit. This is not that surprising taking into account

that the NOON states suffer from the  $2\pi/N$  ambiguity in retrieving the estimated phase, and hence are designed only to work in the local estimation approach when phase fluctuations can be considered small. Derivations of the Heisenberg bounds for phase interferometry using both the QFI and Bayesian approaches are reviewed in Sec. V. We also discuss the problem of deriving the bounds for states with indefinite photon number in which case replacing  $N$  in the derived bounds with the mean number of photons  $\langle N \rangle$  is not always legitimate, so that in some cases the “naive” Heisenberg bound  $1/\langle N \rangle$  may in principle be beaten (Anisimov *et al.*, 2010; Giovannetti and Maccone, 2012; Hofmann, 2009).

Further progress in theoretical quantum metrology stemmed from the need to incorporate realistic decoherence processes in the analysis of the optimal estimation strategies. While deteriorating effects of noise on precision in quantum-enhanced metrological protocols have been realized by many authors working in the field (Caves, 1981; Datta *et al.*, 2011; Gilbert *et al.*, 2008; Huelga *et al.*, 1997; Huver *et al.*, 2008; Rubin and Kaushik, 2007; Sarovar and Milburn, 2006; Shaji and Caves, 2007; Xiao *et al.*, 1987), it has long remained an open question to what extent decoherence effects may be circumvented by employing either more sophisticated states of light or more advanced measurements strategies including e.g. adaptive techniques.

With respect to the most relevant decoherence process in optical applications, i.e. the photonic losses, strong numerical evidence based on the QFI (Demkowicz-Dobrzański *et al.*, 2009; Dörner *et al.*, 2009) indicated that in the asymptotic limit of large number of photons the precision of the optimal quantum protocols approaches  $\text{const}/\sqrt{N}$ , and hence the gain over classical strategies is bound to a constant factor. This fact has been first rigorously proven within the Bayesian approach (Kołodziejczyk and Demkowicz-Dobrzański, 2010) and then independently using the QFI (Knysh *et al.*, 2011). Both approaches yielded the same fundamental bound on precision in the lossy optical interferometry:  $\Delta\varphi \geq \sqrt{(1-\eta)/(\eta N)}$ , where  $\eta$  is the overall power transmission of an interferometric setup. This bound is also valid after replacing  $N$  with  $\langle N \rangle$  when dealing with states of indefinite photon number, and moreover can be easily saturated using the most popular scheme involving a coherent and a squeezed-vacuum state impinging onto two input ports of the Mach-Zehnder interferometer (Caves, 1981). This fact also implies that the presently implemented quantum enhanced schemes in gravitational wave detection, based on interfering the squeezed vacuum with coherent light, operate close to the fundamental bound (Demkowicz-Dobrzański *et al.*, 2013), i.e. they make almost optimal use of non-classical features of light for enhanced sensing given light power and loss levels present in the setup. Based on the mathematical analysis of the geometry of quantum channels (Fujiwara and Imai, 2008;

Matsumoto, 2010) general frameworks have been developed allowing to find fundamental bounds on quantum precision enhancement for general decoherence models (Demkowicz-Dobrzański *et al.*, 2012; Escher *et al.*, 2011). These tools allow to investigate optimality of estimation strategies for basically any decoherence model and typically provide the maximum allowable constant factor improvements forbidding better than  $1/\sqrt{N}$  asymptotic scaling of precision. Detailed presentation of the above mentioned results is given in Sec. VI.

Other approaches to derivation of fundamental metrological bounds have been advocated recently. Making use of the calculus of variations it was shown in (Knysh *et al.*, 2014) how to obtain exact formulas for the achievable asymptotic precision for some decoherence models, while in (Alipour, 2014; Alipour *et al.*, 2014) a variants of QFI have been considered in order to obtain easier to calculate, yet weaker, bounds on precision. While detailed discussion of these approaches is beyond the scope of the present review, in Sec. VI.B.3 we make use of the result from (Knysh *et al.*, 2014) to benchmark the precision bounds derived in the case of phase diffusion noise model.

The paper concludes with Sec. VII with a summary and an outlook on challenging problems in the theory of quantum enhanced metrology.

## II. QUANTUM STATES OF LIGHT

The advent of the laser, light-squeezing and single-photon light sources triggered developments in interferometry that could benefit from the non-classical features of light (Buzek and Knight, 1995; Chekhova, 2011; Torres *et al.*, 2011). In this section, we focus on the quantum-light description of relevance to quantum optical interferometry. We discuss the mode description of light and the most commonly used states in quantum optics—the Gaussian states. In the end, we consider states of definite photon-number and study their particle-description, in particular, investigating their relevant entanglement properties.

### A. Mode description

Classically, electromagnetic field can be divided into orthogonal modes distinguished by their characteristic spatial, temporal and polarization properties. This feature survives in the quantum description of light, where formally we may associate a separate quantum subsystem with each of these modes. Each subsystem is described by its own Hilbert space and, because photons are bosons, can be occupied by an arbitrary number of particles (Mandel and Wolf, 1995; Walls and Milburn, 1995). The most general  $M$ -mode state of light may be

then written as:

$$\rho = \sum_{\mathbf{n}, \mathbf{n}'} \rho_{\mathbf{n}, \mathbf{n}'} |\mathbf{n}\rangle \langle \mathbf{n}'|, \quad \text{Tr}(\rho) = 1, \quad \rho \geq 0 \quad (1)$$

with  $\mathbf{n} = \{n_1, \dots, n_M\}$  and  $|\mathbf{n}\rangle = |n_1\rangle \otimes \dots \otimes |n_M\rangle$  representing a Fock state with *exactly*  $n_i$  photons occupying the  $i$ -th mode. States  $|n_i\rangle$  may be further expressed in terms of the respective creation and annihilation operators  $\hat{a}_i^\dagger, \hat{a}_i$  obeying  $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$ :

$$|n_i\rangle = \frac{\hat{a}_i^{\dagger n_i}}{\sqrt{n_i!}} |0\rangle, \quad \hat{a}_i |n_i\rangle = \sqrt{n_i} |n_i - 1\rangle, \quad \hat{a}_i^\dagger |n_i\rangle = \sqrt{n_i + 1} |n_i + 1\rangle, \quad (2)$$

where  $|0\rangle$  is the vacuum state with no photons at all.

In the context of optical interferometry, modes are typically taken to be distinguishable by their spatial separation, corresponding to different arms of an interferometer, whereas the various optical devices such as mirrors, beam-splitters or phase-delay elements transform the state on its way through the interferometer. Eventually, photon numbers are detected in the output modes allowing to infer the value of the relative phase difference between the arms of the interferometer.

In many applications, the above standard state representation may not be convenient and phase-space description is used instead—in particular, the Wigner function representation (Schleich, 2001; Wigner, 1932). Adopting the convention in which the quadrature operators read  $\hat{x}_i = \hat{a}_i^\dagger + \hat{a}_i$  and  $\hat{p}_i = i(\hat{a}_i^\dagger - \hat{a}_i)$ , the Wigner function may be regarded as a quasi-probability distribution on the quadrature phase space:

$$W(\mathbf{x}, \mathbf{p}) = \frac{1}{(2\pi^2)^M} \int d^M \mathbf{x}' d^M \mathbf{p}' \text{Tr} \left( \rho e^{i[\mathbf{p}'(\hat{\mathbf{x}} - \mathbf{x}) - \mathbf{x}'(\hat{\mathbf{p}} - \mathbf{p})]} \right), \quad (3)$$

where  $\mathbf{x} = \{x_1, \dots, x_M\}$ ,  $\hat{\mathbf{x}} = \{\hat{x}_1, \dots, \hat{x}_M\}$  and similarly for  $\mathbf{p}$  and  $\hat{\mathbf{p}}$ . As a consequence, the Wigner function is real, integrates to 1 over the whole phase space and its marginals yield the correct probability densities of each of the phase space variables. Yet, since it may take negative values it cannot be regarded as a proper probability distribution. Most importantly, it may be reconstructed from experimental data either by tomographic methods (Schleich, 2001) or by direct probing of the phase space (Banaszek *et al.*, 1999), and hence is an extremely useful representation both for theoretical and experimental purposes.

### B. Gaussian states

From the practical point of view, the most interesting class of states are the Gaussian states (D'Ariano *et al.*, 1995; Olivares and Paris, 2007; Pinel *et al.*, 2012, 2013). The great advantage of using them is that they are relatively easy to produce in the laboratory with the help of



standard laser-sources and non-linear optical elements, which allow to introduce non-classical features such as squeezing or entanglement. Gaussian states have found numerous application in various fields of quantum information processing (Adesso, 2006) and are also extensively employed in quantum metrological protocols.

Gaussian states of  $M$  modes are fully characterized by their first and second quadrature moments and are most conveniently represented using the Wigner function which is then just a multidimensional Gaussian distribution

$$W(\mathbf{z}) = \frac{1}{(2\pi)^M \sqrt{\det \sigma}} e^{-\frac{1}{2}(\mathbf{z} - \langle \hat{\mathbf{z}} \rangle)^T \sigma^{-1} (\mathbf{z} - \langle \hat{\mathbf{z}} \rangle)}, \quad (4)$$

where for a more compact notation we have introduced: the phase space variable  $\mathbf{z} = \{x_1, p_1, \dots, x_M, p_M\}$ , the vector containing mean quadrature values  $\langle \hat{\mathbf{z}} \rangle = \{\langle \hat{x}_1 \rangle, \langle \hat{p}_1 \rangle, \dots, \langle \hat{x}_M \rangle, \langle \hat{p}_M \rangle\}$ ,  $\langle \hat{z}_i \rangle = \text{Tr}(\hat{z}_i \rho) = \int d^{2M} \mathbf{z} W(\mathbf{z}) z_i$ , and the  $2M$  dimensional *covariance matrix*  $\sigma$ :

$$\sigma_{i,j} = \frac{1}{2} \langle \hat{z}_i \hat{z}_j + \hat{z}_j \hat{z}_i \rangle - \langle \hat{z}_i \rangle \langle \hat{z}_j \rangle. \quad (5)$$

Gaussian states remain Gaussian under arbitrary evolution involving Hamiltonians at most quadratic in the quadrature operators, what includes all passive devices such as beam-splitters and phase-shifters as well as single- and multi-mode squeezing operations. Below we focus on a few classes of Gaussian states highly relevant to interferometry.

## 1. Coherent states

Coherent states are the Gaussian states with identity covariance matrix  $\sigma = \mathbb{1}$ , so that the uncertainties are equal for all quadratures saturating the Heisenberg uncertainty relations  $\Delta^2 x_i \Delta^2 p_i = 1$  and there are no correlations between the modes. Mean values of quadratures may be arbitrary and correspond to the coherent state complex amplitude  $\boldsymbol{\alpha} = (\langle \hat{\mathbf{x}} \rangle + i \langle \hat{\mathbf{p}} \rangle)/2$ . These are the states produced by any phase-stabilized laser, what makes them almost a fundamental tool in the theoretical description of many quantum optical experiments. Moreover, coherent states have properties that resemble features of classical light, and thus enable to establish a bridge between the quantum and classical descriptions of light.

In the standard representation, an  $M$ -mode coherent state  $|\boldsymbol{\alpha}\rangle = |\alpha_1\rangle \otimes \dots \otimes |\alpha_M\rangle$  is a tensor product of single-mode coherent states, whereas a single-mode coherent state is an eigenstate of the respective annihilation operator:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle, \quad (6)$$

where  $\alpha = |\alpha|e^{i\theta}$  and  $|\alpha|$ ,  $\theta$  are respectively the amplitude and the phase of a coherent state. Equivalently, we may write

$$|\alpha\rangle = \hat{D}(\alpha)|0\rangle, \quad (7)$$

where  $\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}}$  is the so-called *displacement operator* or write the coherent state explicitly as a superposition of consecutive Fock states:

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (8)$$

From the formula above, it is clear that coherent states do not have a definite photon number and if a photon number  $n$  is measured its distribution follows the Poissonian statistics  $P(n) = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}$  with average  $\langle n \rangle = |\alpha|^2$  and standard deviation  $\Delta n = |\alpha|$ . Thus, the relative uncertainty  $\Delta n / \langle n \rangle$  in the measured photon number scales like  $1/\sqrt{\langle n \rangle}$  and hence in the classical limit of large  $\langle n \rangle$  the beam power may be determined up to arbitrary precision. Moreover, the evolution of the coherent state amplitude is identical to the evolution of a classical-wave amplitude. In particular, an optical phase delay  $\varphi$  transforms the state  $|\alpha\rangle$  into  $|\alpha e^{i\varphi}\rangle$ . These facts justify a common jargon of calling coherent states the classical states of light, even though for relatively small amplitudes different coherent states may be hard to distinguish due to their non-orthogonality  $|\langle \alpha | \beta \rangle|^2 = e^{-|\alpha - \beta|^2}$ . More generally, we call  $\rho_{\text{cl}}$  a classical state of light if and only if it can be written as a mixture of coherent states:

$$\rho_{\text{cl}} = \int d^{2M} \boldsymbol{\alpha} P(\boldsymbol{\alpha}) |\boldsymbol{\alpha}\rangle \langle \boldsymbol{\alpha}| \quad (9)$$

with  $P(\boldsymbol{\alpha}) \geq 0$ , which is equivalent to the statement that  $\rho_{\text{cl}}$  admits a non-negative Glauber  $P$ -representation (Glauber, 1963; Walls and Milburn, 1995). Classical states are often used as a benchmark to test the degree of possible quantum enhancement which may be obtained by using more general states outside this class.

## 2. Single-mode squeezed states

Heisenberg uncertainty principle imposes that  $\Delta x \Delta p \geq 1$  for all possible quantum states. Single-mode states that saturate this inequality are called the *single-mode squeezed states* (Walls and Milburn, 1995). As mentioned above, coherent states fall into such a category serving as a special example for which  $\Delta x = \Delta p = 1$ . Yet, as for general squeezed states  $\Delta x \neq \Delta p$ , the noise in one of the quadratures can be made smaller than in the other. Formally, a single-mode squeezed state may always be expressed as

$$|\alpha, r\rangle = \hat{D}(\alpha) \hat{S}(r) |0\rangle, \quad (10)$$

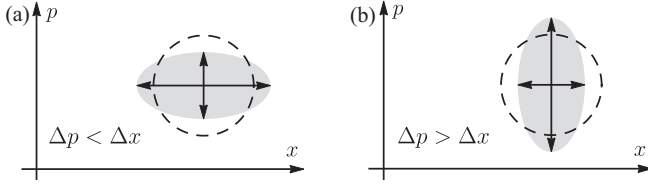


FIG. 1 Phase-space diagrams denoting uncertainties in different quadratures for momentum squeezed states (a) and for position squeezed states (b). Dashed circles denote corresponding uncertainties for coherent states  $\Delta x = \Delta p$ .

where  $\hat{S}(r) = \exp(\frac{1}{2}r^*\hat{a}^2 - \frac{1}{2}r\hat{a}^{\dagger 2})$  is the *squeezing operator*,  $r = |r|e^{i\theta}$  is a complex number and  $|r|$  and  $\theta$  are the *squeezing factor* and the *squeezing angle* respectively. In fact, any pure Gaussian one-mode state may be written in the above form. For  $\theta = 0$ , uncertainties in the quadratures  $x$  and  $p$  read  $\Delta x = e^{-r}$  and  $\Delta p = e^r$ —reduction of noise in one quadrature is accompanied by an added noise in the other one. This may be conveniently visualized in the phase-space picture by error disks representing uncertainty in quadratures in different directions, see Fig. 1. In such a representation squeezed states correspond to ellipses while coherent states are represented by circles. Importantly, the fact that the uncertainty of one of the quadratures can be less relatively to the other makes it possible to design an interferometric scheme where the measured photon number fluctuations are below that of a coherent state and allows for a sub-shot noise phase estimation precision, see Sec. III. As squeezed states in general cannot be described as mixtures of coherent states, they are non-classical and their features cannot be fully described by classical electrodynamics. Nevertheless, they can be relatively easily prepared using non-linear optical elements in the process of parametric down conversion (Bachor and Ralph, 2004).

The special type of squeezed states which is most relevant from the metrological perspective is the class of the *squeezed vacuum states* that possess vanishing mean values of their quadratures, i.e.  $\langle \hat{z} \rangle = 0$ :

$$|r\rangle = \hat{S}(r)|0\rangle. \quad (11)$$

In the Fock basis a squeezed vacuum state reads

$$|r\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{H_n(0)}{\sqrt{n!}} \left(\frac{\tanh r}{2}\right)^{\frac{n}{2}} e^{i\frac{n\theta}{2}} |n\rangle, \quad (12)$$

where  $H_n(0)$  denotes values of  $n$ -th Hermite polynomial at  $x = 0$ . As for odd  $n$   $H_n(x)$  is antisymmetric and thus  $H_n(0) = 0$ , it follows that squeezed vacuum states are superpositions of Fock states with only *even* photon numbers. The average number of photons in a squeezed vacuum state is given by  $\langle n \rangle = \sinh^2 r$ , what means that, despite their name, a squeezed vacuum states contain photons, possibly a lot of them.

### 3. Two-mode squeezed states

The simplest non-classical two-mode Gaussian state is the so-called two-mode squeezed vacuum state or the *twin-beam state* (Walls and Milburn, 1995). Mathematically, such a state is generated from the vacuum by a two-mode squeezing operation, so that

$$|\xi\rangle_2 = \hat{S}_2(\xi)|0,0\rangle, \quad (13)$$

where  $\hat{S}_2(\xi) = \exp(\xi^* \hat{a} \hat{b} - \xi \hat{a}^\dagger \hat{b}^\dagger)$  and  $\xi = |\xi|e^{i\theta}$ , whereas in the Fock basis it reads

$$|\xi\rangle_2 = \frac{1}{\cosh \xi} \sum_{n=0}^{\infty} (-1)^n e^{i\theta} \tanh^n \xi |n,n\rangle. \quad (14)$$

A notable feature of the twin-beam state, which may be clearly seen from Eq. (14), is that it is not a product of squeezed states in modes  $a$  and  $b$ , but rather it is correlated in between them being a superposition of terms with the same number of photons in both modes. Its first moments of all quadratures are zero,  $\langle \hat{z} \rangle = 0$ , whereas in the case of  $\xi = |\xi|$  its covariance matrix has a particularly simple form:

$$\sigma = \begin{pmatrix} \cosh(2\xi) & 0 & \sinh(2\xi) & 0 \\ 0 & \cosh(2\xi) & 0 & -\sinh(2\xi) \\ \sinh(2\xi) & 0 & \cosh(2\xi) & 0 \\ 0 & -\sinh(2\xi) & 0 & \cosh(2\xi) \end{pmatrix}. \quad (15)$$

Such a covariance matrix clearly indicates the presence of correlations between the modes and since the state (13) is pure this implies immediately the presence of the *mode-entanglement*. In fact, in the limit of large squeezing coefficient  $|\xi| \rightarrow \infty$  such twin-beam state becomes the original famous Einstein-Podolsky-Rosen state (Adesso, 2006; Banaszek and Wódkiewicz, 1998) that violates assumptions of any realistic local hidden variable theory.

Twin-beam states may be generated in a laboratory by various non-linear processes such as four- and three-wave mixing (Bachor and Ralph, 2004; Reid and Drummond, 1988). Alternatively, they may be produced by mixing two single-mode squeezed vacuum states with opposite squeezing angles on a fifty-fifty beam-splitter.

### 4. General two-mode Gaussian states

General two-mode Gaussian state is rather difficult to write in the Fock basis, so it is best characterized by its  $4 \times 4$  real symmetric covariance matrix

$$\sigma = \begin{pmatrix} [\sigma_{11}] & [\sigma_{12}] \\ [\sigma_{21}] & [\sigma_{22}] \end{pmatrix}, \quad (16)$$

where  $\sigma_{ij}$  represent blocks with  $2 \times 2$  matrices describing correlations between the  $i$ -th and the  $j$ -th mode, and the vector of the first moments  $\langle \hat{z} \rangle =$

$\{\langle \hat{x}_1 \rangle, \langle \hat{p}_1 \rangle, \langle \hat{x}_2 \rangle, \langle \hat{p}_2 \rangle\}$ . This in total gives up to fourteen real parameters describing the state: ten covariances, two displacement amplitudes and two phases of displacement. General Gaussian states are in principle feasible within current technological state of art, as any pure Gaussian state can theoretically be generated from the vacuum by utilizing only a combination of one-, two-mode squeezing and displacement operations with help of beam-splitters and one-mode rotations (Adesso, 2006). Furthermore, mixed Gaussian states are obtained as a result of tracing out some of the system degrees of freedom, which is effectively the case in the presence of light losses, or by adding a Gaussian noise to the state.

### C. Definite photon number states

Gaussian states are important from the practical point of view due to the relative ease with which they may be prepared. From a conceptual point of view, however, when asking fundamental questions on limits to quantum enhancement in interferometry, states with a definite photon number prove to be a better choice. The main reason is that photons are typically regarded as a resource in interferometry and when benchmarking different interferometric schemes it is natural to restrict the class of states with the same number of photons, i.e. the same resources consumed. A general  $M$ -mode state consisting of  $N$  photons is given by:

$$\rho_N = \sum_{|\mathbf{n}|=|\mathbf{n}'|=N} \rho_{\mathbf{n},\mathbf{n}'} |\mathbf{n}\rangle \langle \mathbf{n}'|, \quad (17)$$

where  $|\mathbf{n}| = \sum_i n_i$ , so that the summation is restricted only to terms with exactly  $N$  photons in all the modes. Apart from the vacuum state  $|0\rangle$  no Gaussian state falls into this category.

States with an exact photon number are extensively used in other fields of quantum information processing, including quantum communication and quantum computing (Kok *et al.*, 2007; Pan *et al.*, 2012). Most of the quantum computation and communication schemes are designed with such states in mind, as they provide the most intuitive and clear picture of the role the quantum features play in these tasks. For large  $N$ , however, states with a definite photon numbers are notoriously hard to prepare and states with  $N$  of the order of 10 can only be produced with the present technology pushed to its limits (Hofheinz *et al.*, 2008; Sayrin *et al.*, 2011; Torres *et al.*, 2011). When considering states with definite  $N$ , it is also possible to easily switch between the mode- and particle-description of the states of light, which is a feature that we discuss in the following section.

### D. Particle description

When dealing with states of definite photon number, instead of thinking about modes as quantum subsystems that possess some number of excitations (photons), we may equivalently consider the “first quantization” formalism and regard photons themselves as elementary subsystems. Fundamentally, photons are indistinguishable particles and since they are bosons their wave function should always be permutation-symmetric. Still, it is common in the literature to use a description in which photons are regarded as distinguishable particles and adopt a notation such that

$$|\mathbf{m}\rangle = |m_1\rangle_1 \otimes |m_2\rangle_2 \otimes \cdots \otimes |m_N\rangle_N \quad (18)$$

denotes a product state of  $N$  photons, where the  $i$ -th photon occupies the mode  $m_i$ . We explicitly add subscripts to the kets above labeling each constituent photon, in order to distinguish this notation from the mode description of Eq. (1), where kets denoted various modes and not the distinct particles. The description (18) is legitimate provided there are some degrees of freedom that ascribe a meaning to the statement “the  $i$ -th photon”. For example, in the case when photons are prepared in non-overlapping time-bins, the time-bin degree of freedom plays the role of the label indicating a *particular* photon, whereas the spatial characteristics determine the state of a *given* photon. Nevertheless, if we assume the overall wave function describing also the temporal degrees of freedom of the complete state to be fully symmetric, the notion of the “ $i$ -th photon” becomes meaningless.

A general pure state of  $N$  “distinguishable” photons has the form:

$$|\psi_N\rangle = \sum_{\mathbf{m}} c_{\mathbf{m}} |\mathbf{m}\rangle \quad (19)$$

where  $\sum_{\mathbf{m}} |c_{\mathbf{m}}|^2 = 1$ . If indeed there is no additional degree of freedom that makes the notion of “the  $i$ -th photon” meaningful, the above state should possess the symmetry property such that  $c_{\mathbf{m}} = c_{\Pi(\mathbf{m})}$ , where  $\Pi$  is an arbitrary permutation of the  $N$  indices.

Consider for example a Fock state  $|\mathbf{n}\rangle = |n_1\rangle \dots |n_M\rangle$  of  $N = n_1 + \dots + n_M$  indistinguishable photons in  $M$  modes. In the particle description the state has the form:

$$|\mathbf{n}\rangle = \sqrt{\frac{n_1! \dots n_M!}{N!}} \times \sum_{\Pi} |\Pi(\underbrace{1, \dots, 1}_{n_1}, \underbrace{2, \dots, 2}_{n_2}, \dots, \underbrace{M, \dots, M}_{n_M})\rangle, \quad (20)$$

where the sum is performed over all non-trivial permutations  $\Pi$  of the indices inside the curly brackets (Shankar, 1994). Since all quantum states may be written in the Fock basis representation, by the above construction one can always translate any quantum state to the particle description.

### E. Mode vs particle entanglement

One of the most important features which makes the quantum theory different from the classical one is the notion of entanglement (Horodecki *et al.*, 2009). This phenomenon plays also an important role in quantum metrology and is often claimed to be the crucial resource for the enhancement of the measurement precision (Giovannetti *et al.*, 2006; Pezzé and Smerzi, 2009). Conflicting statements can be found in the literature, however, as some of the authors claim that entanglement is not indispensable to get a quantum precision enhancement (Benatti and Braun, 2013). This confusion stems simply from the fact that entanglement is a relative concept dependent on the way we divide the relevant Hilbert space into particular subsystems. In order to clarify these issues, it is necessary to explicitly study relation between mode- and particle-entanglement, i.e. entanglement with respect to different tensor product structures used in the two descriptions.

Firstly, let us go through basic definitions and notions of entanglement. The state  $\rho_{AB}$  of two parties  $A$  and  $B$  is called separable if and only if one can write it as a mixture of product of states of individual subsystems:

$$\rho^{(AB)} = \sum_i p_i \rho_i^{(A)} \otimes \rho_i^{(B)}, \quad p_i \geq 0. \quad (21)$$

Entangled states are defined as all states that are not separable. A crucial feature of entanglement is that it depends on the *division* into subsystems. For example, consider three qubits  $A$ ,  $B$  and  $C$  and their joint quantum state  $\rho^{(ABC)} = \sum_{i,j=0}^1 \frac{1}{2} |i\rangle\langle j| \otimes |i\rangle\langle j| \otimes |0\rangle\langle 0|$ . This state is separable with respect to the  $AB|C$  cut but is entangled with respect to the  $A|BC$  cut.

As a first example, consider a two-mode Fock state  $|1\rangle_a |1\rangle_b$  which represents one photon in mode  $a$  and one photon in mode  $b$ . This state written in the mode formalism of Eq. (1) is clearly separable. On the other hand, photons are indistinguishable bosons and if we would like to write their state in the particle formalism of Eq. (18) we have to symmetrize over all possible permutations of particles, thus obtaining the state

$$|1\rangle_a |1\rangle_b = \frac{1}{\sqrt{2}} (|a\rangle_1 |b\rangle_2 + |b\rangle_1 |a\rangle_2). \quad (22)$$

In this representation the state is clearly entangled. We may thus say that the state contains *particle* entanglement but *not* the *mode* entanglement.

If, however, we perform the Hong-Ou-Mandel experiment and send the  $|1\rangle_a |1\rangle_b$  state through a balanced beam-splitter which transforms mode annihilation operators as  $\hat{a} \rightarrow (\hat{a} + \hat{b})/\sqrt{2}$ ,  $\hat{b} \rightarrow (\hat{a} - \hat{b})/\sqrt{2}$ , the resulting state reads:

$$|1\rangle_a |1\rangle_b \rightarrow \frac{1}{\sqrt{2}} (|2\rangle_a |0\rangle_b - |0\rangle_a |2\rangle_b) = \frac{1}{\sqrt{2}} (|a\rangle_1 |a\rangle_2 - |b\rangle_1 |b\rangle_2), \quad (23)$$

which is *both* mode- and particle-entangled. Mode entanglement emerges because the beam-splitter is a joint operation over two modes that introduces correlations between them. On the other hand, it is a local operation with respect to the particles, i.e. it can be written as  $U \otimes U$  in the particle representation, and does not couple photons with each other. Thus, using a beam-splitter one may change mode entanglement but not the content of particle entanglement.

As a second example, consider two modes of light,  $a$  and  $b$ , each of them in coherent state with the same amplitude  $\alpha$ ,  $|\alpha\rangle_a |\alpha\rangle_b$ . This state clearly has *no* mode entanglement. Since this state does not have a definite photon number, in order to ask questions about the particle entanglement we first need to consider its projection on one of the  $N$ -photon subspaces—one can think of a non-demolition total photon-number measurement yielding result  $N$ . After normalizing the projected state we obtain:

$$|\psi_N\rangle = [|\alpha\rangle_a |\alpha\rangle_b]^{(N)} = \frac{1}{\sqrt{2^N}} \sum_{n=0}^N \sqrt{\binom{N}{n}} |n\rangle_a |N-n\rangle_b, \quad (24)$$

which in the particle representation reads:

$$|\psi_N\rangle = \frac{1}{\sqrt{2^N}} \bigotimes_{i=1}^N (|a\rangle_i + |b\rangle_i) \quad (25)$$

and is clearly a separable state. The fact that products of coherent states contain no particle entanglement is in agreement with our definition of classical state given in Eq. (9) being a mixture of products of coherent states. A *classical state* according to this definition will contain *neither* mode nor particle entanglement.

As a last example, consider the case of particular interest for quantum interferometry, i.e. a coherent state of mode  $a$  and a squeezed vacuum state of mode  $b$ :  $|\alpha\rangle_a |r\rangle_b$ . Again this state has *no* mode entanglement. On the other hand it *is* particle entangled. To see this, consider e.g. the two-photon sector, which up to irrelevant normalization factor reads:

$$[|\alpha\rangle |r\rangle]^{(N=2)} \propto \alpha^2 |2\rangle_a |0\rangle_b + \tanh r |0\rangle_a |2\rangle_b = \alpha^2 |a\rangle_1 |a\rangle_2 + \tanh r |b\rangle_1 |b\rangle_2 \quad (26)$$

and contains particle entanglement provided both  $\alpha$  and  $r$  are non-zero. We argue and give detailed arguments in Sec. V that it is indeed the particle entanglement and *not* the mode entanglement that is relevant in quantum-enhanced interferometry. See also Killoran *et al.* (2014) for more insight into the relation between mode and particle entanglement.

### III. MACH-ZEHNDER INTERFEROMETRY

We begin the discussion of quantum-enhancement effects in optical interferometry by discussing the paradigm



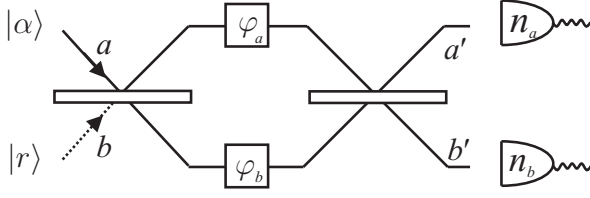


FIG. 2 The Mach-Zehnder interferometer, with two input light modes  $a, b$  and two output modes  $a', b'$ . In a standard configuration a coherent state of light  $|\alpha\rangle$  is sent into mode  $a$ . In order to obtain quantum enhancement, one needs to make use of the  $b$  input port also, sending e.g. the squeezed vacuum state  $|r\rangle$ .

matic model of the Mach-Zehnder (MZ) interferometer. We analyze the most popular interferometric schemes involving the use of coherent and squeezed states of light accompanied by a basic measurement-estimation procedure, in which the phase is estimated based on the value of the photon-number difference between the two output ports of the interferometer. Such a protocol provides us with a benchmark that we may use in the following sections when discussing the optimality of the interferometry schemes both with respect to the states of light used as well as the measurements and the estimation procedures employed.

### A. Phase-sensing uncertainty

In the standard MZ configuration, depicted in Fig. 2, a coherent state of light is split on a balanced beam-splitter, the two beams acquire phases  $\varphi_a, \varphi_b$  respectively, interfere on the second beam-splitter and finally the photon numbers  $n_a, n_b$  are measured at the output ports. Let  $\hat{a}, \hat{b}$  and  $\hat{a}', \hat{b}'$  be the annihilation operators corresponding to the two input and the two output modes respectively. The combined action of the beam-splitters and the phase delays results in the effective transformation of the annihilation operators:

$$\begin{pmatrix} \hat{a}' \\ \hat{b}' \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{i\varphi_a} & 0 \\ 0 & e^{i\varphi_b} \end{pmatrix} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = e^{i(\varphi_a + \varphi_b)/2} \begin{pmatrix} \cos(\varphi/2) & -\sin(\varphi/2) \\ \sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \quad (27)$$

where  $\varphi = \varphi_b - \varphi_a$  is the relative phase delay and for convenience we assume that the beams acquire a  $-\pi/2$  or  $\pi/2$  phase when transmitted through the first or the second beam-splitter respectively. The common phase factor  $e^{i(\varphi_a + \varphi_b)/2}$  is irrelevant for further discussion in this section and will be omitted.

In order to get a better insight into the quantum-enhancement effects in the operation of the MZ interferometer, it is useful to make use of the so-called Jordan-Schwinger map (Schwinger, 1965) and analyse the action

of the MZ interferometer in terms of the algebra of the angular momentum operators (Yurke *et al.*, 1986). Let us define the operators:

$$\hat{J}_x = \frac{1}{2}(\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a}), \quad \hat{J}_y = \frac{i}{2}(\hat{b}^\dagger \hat{a} - \hat{a}^\dagger \hat{b}), \quad \hat{J}_z = \frac{1}{2}(\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b}), \quad (28)$$

which fulfill the angular momentum commutation relations  $[\hat{J}_i, \hat{J}_j] = i\epsilon_{ijk}\hat{J}_k$  while the corresponding square of the total angular momentum reads:

$$\hat{J}^2 = \frac{\hat{N}}{2} \left( \frac{\hat{N}}{2} + 1 \right), \quad \hat{N} = \hat{a}^\dagger \hat{a} + \hat{b}^\dagger \hat{b}, \quad (29)$$

where  $\hat{N}$  is the total photon number operator. The action of linear optical elements appearing in the MZ interferometer can now be described as rotations in the abstract spin space:  $\hat{a}' = U\hat{a}U^\dagger, \hat{b}' = U\hat{b}U^\dagger, U = \exp(-i\alpha\hat{\mathbf{J}} \cdot \mathbf{s})$ , where  $\hat{\mathbf{J}} = \{\hat{J}_x, \hat{J}_y, \hat{J}_z\}$  and  $\alpha, \mathbf{s}$  are the angle and the axis of the rotation respectively. In particular, the balanced beam splitter is a rotation around the  $x$  axis by an angle  $\pi/2$ :  $U = \exp(-i\frac{\pi}{2}\hat{J}_x)$ , while the phase delay is a  $\varphi$  rotation around the  $z$  axis:  $U = \exp(-i\varphi\hat{J}_z)$ . Instead of analysing the transformation of the annihilation operators, it is more convenient to look at the corresponding transformation of the  $J_i$  operators themselves:

$$\begin{pmatrix} \hat{J}'_x \\ \hat{J}'_y \\ \hat{J}'_z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{J}_x \\ \hat{J}_y \\ \hat{J}_z \end{pmatrix} = \begin{pmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{pmatrix} \begin{pmatrix} \hat{J}_x \\ \hat{J}_y \\ \hat{J}_z \end{pmatrix}, \quad (30)$$

which makes it clear that the sequence of  $\pi/2, \varphi$  and  $-\pi/2$  rotations around axes  $x, z$  and  $x$  respectively, results in an effective  $\varphi$  rotation around the  $y$  axis.

Using the above formalism, let us now derive a simple formula for uncertainty of phase-sensing based on the measurement of the photon-number difference at the output. Note that  $\hat{n}_a - \hat{n}_b = 2\hat{J}_z$ , so the photon-number difference measurement is equivalent to the  $\hat{J}_z$  measurement. Utilizing Eq. (30) in the Heisenberg picture, the average  $J_z$  evaluated on the interferometer output state may be related to the average of  $J'_z$  of the input state  $|\psi\rangle_{\text{in}}$  as

$$\langle \hat{J}_z \rangle = \cos \varphi \langle \hat{J}_z \rangle_{\text{in}} - \sin \varphi \langle \hat{J}_x \rangle_{\text{in}}. \quad (31)$$

In order to assess the precision of  $\varphi$ -estimation, we also calculate the variance of the  $\hat{J}_z$  operator of the output state of the interferometer:

$$\Delta^2 J_z = \cos^2 \varphi \Delta^2 J_z|_{\text{in}} + \sin^2 \varphi \Delta^2 J_x|_{\text{in}} + 2 \sin \varphi \cos \varphi \text{cov}(J_x, J_z)|_{\text{in}}, \quad (32)$$

where  $\text{cov}(J_x, J_z) = \frac{1}{2}\langle \hat{J}_x \hat{J}_z + \hat{J}_z \hat{J}_x \rangle - \langle \hat{J}_x \rangle \langle \hat{J}_z \rangle$  is the covariance of the two observables. The precision of estimating  $\varphi$  can now be quantified via a simple error-propagation formula:

$$\Delta\varphi = \frac{\Delta J_z}{\left| \frac{d\langle \hat{J}_z \rangle}{d\varphi} \right|}. \quad (33)$$

### B. Coherent-state interferometry

Let us now analyze the precision given by Eq. (33) for the standard optical interferometry with the input state  $|\psi\rangle_{\text{in}} = |\alpha\rangle|0\rangle$ , representing a coherent state and no light at all being sent into the input modes  $a$  and  $b$  of the interferometer in Fig. 2. The relevant quantities required for calculating the precision given in Eq. (33) read:

$$\begin{aligned} \langle \hat{J}_z \rangle_{\text{in}} &= \frac{1}{2}|\alpha|^2, \quad \langle \hat{J}_x \rangle_{\text{in}} = 0, \quad \Delta^2 J_z|_{\text{in}} = \Delta^2 J_x|_{\text{in}} = \frac{1}{4}|\alpha|^2, \\ \text{cov}(J_x, J_z)|_{\text{in}} &= 0 \end{aligned} \quad (34)$$

yielding the precision:

$$\Delta\varphi^{|\alpha\rangle|0\rangle} = \frac{\frac{1}{2}|\alpha|}{\frac{1}{2}|\alpha|^2|\sin\varphi|} = \frac{1}{|\alpha\sin\varphi|} = \frac{1}{\sqrt{\langle N \rangle}|\sin\varphi|}, \quad (35)$$

where the average photon number  $\langle N \rangle = \langle \hat{N} \rangle = |\alpha|^2$ . The above formula represents  $1/\sqrt{\langle N \rangle}$  shot noise scaling of precision characteristic for the classical interferometry. The shot noise is a consequence of the  $\Delta J_z$  effectively representing the Poissonian fluctuations of the photon-number difference measurements at the output ports. Yet, although such fluctuations are  $\varphi$ -independent, the average photon-number difference  $\langle J_z \rangle$  changes with  $\varphi$  with speed proportional to  $|\sin\varphi|$  appearing in Eq. (35), so that the optimal operating points are at  $\varphi = \pi/2, 3\pi/2$ .

### C. Fock state interferometry

We can attempt to reduce the estimation uncertainty using more general states of light at the input. For example, we can replace the coherent state with an  $N$ -photon Fock state, so that  $|\psi\rangle_{\text{in}} = |N\rangle|0\rangle$ . This is an eigenstate of  $\hat{J}_z$  and hence  $\Delta^2 J_z|_{\text{in}} = 0$ , and only the  $\Delta^2 J_x|_{\text{in}}$  contributes to the  $\hat{J}_z$  variance at the output:  $\Delta^2 J_z = \frac{1}{4}N\sin^2\varphi$ . Since  $\langle \hat{J}_z \rangle = \frac{1}{2}N\cos\varphi$ , the corresponding estimation uncertainty reads:

$$\Delta\varphi^{|N\rangle|0\rangle} = \frac{1}{\sqrt{N}}, \quad (36)$$

being again shot-noise limited. The sole benefit of using the Fock state is the lack of  $\varphi$ -dependence of the

estimation precision. This, however, is scarcely of any use in practice, since one may always perform rough interferometric measurements and bring the setup close to the optimal operating points before performing more precise measurements there. Moreover, the  $\varphi$ -dependence in Eq. (35) can be easily removed by taking into account not only the photon-number difference observable  $\hat{J}_z$  but also the total photon number measured. By using their ratio as an effective observable in the r.h.s. of Eq. (33), or in other words by considering the “visibility observable”, the formula (35) is replaced by 36.

### D. Coherent + squeezed-vacuum interferometry

We thus need to use more general input states in order to surpass the shot noise limit. Firstly, let us note that sending the light solely to one of the input ports will not provide us with the desired benefit. As in the end only a photon-number measurement is assumed, which is not sensitive to any relative phase differences between various Fock terms of the output state, any scenario involving a single-beam input may always be translated to the situation in which an incoherent mixture of Fock states is sent onto the input port. Since the variance is a convex function with respect to state density matrices, and thus increasing under mixing, such strategies are of no use for our purposes.

Let us now consider a scheme were apart from the coherent light we additionally send a squeezed-vacuum state into the other input port (Caves, 1981):  $|\psi\rangle_{\text{in}} = |\alpha\rangle \otimes |r\rangle$ , see Fig. 2. This kind of strategy is being implemented in current most advanced interferometers designed to detect gravitational waves like LIGO or GEO600 (LIGO Collaboration, 2011, 2013; Pitkin *et al.*, 2011). Assuming for simplicity that  $r$  is real, the relevant quantities required to calculate the estimation precision read:

$$\begin{aligned} \langle \hat{N} \rangle &= |\alpha|^2 + \sinh^2 r, \quad \langle J_z \rangle_{\text{in}} = \frac{1}{2}(|\alpha|^2 - \sinh^2 r), \quad \langle J_x \rangle_{\text{in}} = 0 \\ \Delta^2 J_z|_{\text{in}} &= \frac{1}{4}(|\alpha|^2 + \frac{1}{2}\sinh^2 2r), \quad \text{cov}(J_x, J_z)|_{\text{in}} = 0, \\ \Delta^2 J_x|_{\text{in}} &= \frac{1}{4}(|\alpha|^2 \cosh 2r - \text{Re}(\alpha^2) \sinh 2r + \sinh^2 r). \end{aligned} \quad (37)$$

Hence, the usage of squeezed-vacuum as a second input allows to reduce the variance  $\Delta^2 J_x|_{\text{in}}$  thanks to the  $\text{Re}(\alpha^2) \sinh 2r$  term above, which is then maximized by choosing the phase of the coherent state such that  $\alpha = \text{Re}(\alpha)$ . This corresponds to the situation, when the coherent state is displaced in phase space in the direction in which the squeezed vacuum possesses its lowest variance. With such an optimal choice of phase, substituting the above formulas into Eq. (33), we obtain the

final expression for the phase-estimation precision:

$$\Delta\varphi^{|\alpha\rangle|r\rangle} = \frac{\sqrt{\cot^2\varphi(|\alpha|^2 + \frac{1}{2}\sinh^2 2r) + |\alpha|^2 e^{-2r} + \sinh^2 r}}{||\alpha|^2 - \sinh^2 r|}. \quad (38)$$

The optimal operation point are again clearly  $\varphi = \pi/2, 3\pi/2$ , since at them the first term under the square root, which is non-negative, vanishes. For a fair comparison with other strategies we should fix the total average number of photons  $\langle N \rangle$ , which is regarded as a resource, and optimize the split of energy between the coherent and the squeezed vacuum beams in order to minimize  $\Delta\varphi$ . This optimization can only be done numerically, but the solution can be well approximated analytically in the regime of  $\langle N \rangle \gg 1$ . In this regime the squeezed vacuum should carry approximately  $\sqrt{\langle N \rangle}/2$  of photons, so the squeezing factor obeys  $\sinh^2 r \approx \frac{1}{4}e^{2r} \approx \sqrt{\langle N \rangle}/2$  while the majority of photons belongs to the coherent beam. The resulting precision reads

$$\Delta\varphi^{|\alpha\rangle|r\rangle} \stackrel{\langle N \rangle \gg 1}{\approx} \frac{\sqrt{\langle N \rangle/(2\sqrt{\langle N \rangle})} + \sqrt{\langle N \rangle}/2}{\langle N \rangle - \sqrt{\langle N \rangle}/2} \stackrel{\langle N \rangle \gg 1}{\approx} \frac{1}{\langle N \rangle^{3/4}} \quad (39)$$

and proves that indeed this strategy offers better than shot noise scaling of precision.

While the above example shows that indeed quantum states of light may lead to an improved sensitivity, the issue of optimality of the proposed scheme has not been addressed. In fact, keeping the measurement-estimation scheme unchanged, it is possible to further reduce the estimation uncertainty by sending a more general two-mode Gaussian states of light with squeezing present in both input ports and reach the  $\propto 1/\langle N \rangle$  scaling of precision (Olivares and Paris, 2007; Yurke *et al.*, 1986). We skip the details here, as the optimization of the general Gaussian two-mode input state minimizing the estimation uncertainty of the scheme considered is cumbersome. More importantly, using the tools of estimation theory introduced in Sec. IV, we will later show in Sec. V that even with  $|\alpha\rangle \otimes |r\rangle$  class of input states it is possible to reach the  $\propto 1/\langle N \rangle$  scaling of precision, but this requires a significantly modified measurement-estimation scheme (Pezzé and Smerzi, 2008) and different energy partition between the two input modes.

### E. Definite photon-number state interferometry

Even though definite photon-number states are technically difficult to prepare, they are conceptually appealing and we make use of them to demonstrate explicitly the possibility of achieving the  $1/N$  Heisenberg scaling of estimation precision in Mach-Zehnder interferometry. We have already shown that sending an  $N$ -photon state into a single input port of the interferometer does not lead to

an improved precision compared with a coherent-state-based strategy. Therefore, we need to consider states with photons being simultaneously sent into both input ports. A general  $N$ -photon two-mode input state can be written down using the angular momentum notation as

$$|\psi\rangle_{\text{in}} = \sum_{m=-j}^j c_m |j, m\rangle, \quad (40)$$

where  $j = N/2$  and  $|j, m\rangle = |j+m\rangle|j-m\rangle$  in the standard mode-occupation notation. In particular,  $|j, j\rangle = |N\rangle|0\rangle$  corresponds to a state with all the photons being sent into the upper input port. One can easily check that angular momentum operators introduced in Eq. (28) act in a standard way on the  $|j, m\rangle$  states. For concreteness assume  $N$  is even and consider the state (Yurke *et al.*, 1986):

$$\begin{aligned} |\psi\rangle_{\text{in}} &= \frac{1}{\sqrt{2}} (|j, 0\rangle + |j, 1\rangle) = \\ &= \frac{1}{\sqrt{2}} (|\frac{N}{2}\rangle|\frac{N}{2}\rangle + |\frac{N}{2}+1\rangle|\frac{N}{2}-1\rangle), \end{aligned} \quad (41)$$

for which:

$$\begin{aligned} \langle \hat{J}_z \rangle_{\text{in}} &= \frac{1}{2}, \quad \langle \hat{J}_x \rangle_{\text{in}} = \frac{1}{2}\sqrt{j(j+1)}, \quad \Delta^2 J_z|_{\text{in}} = \frac{1}{2}, \\ \Delta^2 J_x|_{\text{in}} &= \frac{1}{2}j(j+1) - \frac{1}{4}, \quad \text{cov}(J_x, J_z)|_{\text{in}} = 0. \end{aligned} \quad (42)$$

Plugging the above expressions into Eq. (33), we get:

$$\Delta\varphi = \frac{\sqrt{\cos^2\varphi + \sin^2\varphi[j(j+1) - 1]}}{|\sin\varphi + \cos\varphi\sqrt{j(j+1)}|}. \quad (43)$$

The optimal operation point corresponds to  $\sin\varphi = 0$ , where we benefit from large  $\langle \hat{J}_x \rangle_{\text{in}}$  and low  $\Delta^2 J_z|_{\text{in}}$  making the state very sensitive to rotations around the  $y$  axis. The resulting precision reads:

$$\Delta\varphi = \frac{1}{\sqrt{j(j+1)}} \stackrel{N \gg 1}{\approx} \frac{2}{N}, \quad (44)$$

indicating the possibility of achieving the Heisenberg scaling of precision. We should note here, that a simpler state  $|\psi\rangle_{\text{in}} = |j, 0\rangle = |N/2\rangle|N/2\rangle$  called the *twin-Fock state* where  $N/2$  photons are simultaneously sent into each of the input ports, is also capable of providing the Heisenberg scaling of precision (Holland and Burnett, 1993), but requires a different measurement-estimation scheme which goes beyond the analysis of the average photon-number difference at the output, see Sec. V.C.3.

One should bear in mind that the expressions for attainable precision presented in this section base on a simple error propagation formula calculated at a particular operating point. Therefore, in order to approach any of

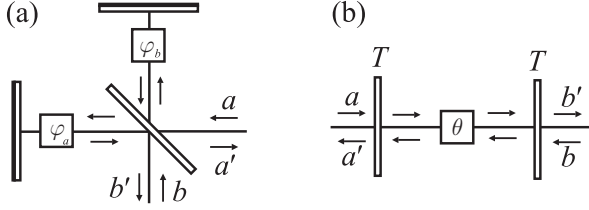


FIG. 3 Other popular two-input/two-output mode interferometers: (a) Michelson interferometer with one-way phase delays  $\varphi_a$ ,  $\varphi_b$  in the respective arms, (b) Fabry-Pérot interferometer with one-way phase delay  $\theta$  and power transmission  $T$  of the mirrors.

the precisions claimed, one needs to first lock the interferometer to operate close to an optimal point, which also requires some of the resources to be consumed. Rigorous quantification of the total resources needed to attain a given estimation precision starting with a completely unknown phase may be difficult in general. We return to this issue in Sec. V, where we are able to resolve this problem by approaching it with the language of the Bayesian inference.

### F. Other interferometers

Even though we have focused on the Mach-Zehnder interferometer setup, analogous results could be obtained for other optical interferometric configurations, such as the Michelson interferometer, the Fabry-Pérot interferometer, as well as the atomic interferometry setups utilized in: atomic clocks operation (Diddams *et al.*, 2004), spectroscopy (Leibfried *et al.*, 2004), magnetometry (Budker and Romalis, 2007) or the BEC interferometry (Cronin *et al.*, 2009). We briefly show below that despite physical differences the mathematical framework is common to all these cases and as such the results presented in this review, even though derived with the simple optical interferometry in mind, have a much broader scope of applicability.

The Michelson interferometer is depicted in Fig. 3a. Provided the output modes  $a'$ ,  $b'$  can be separated from the input ones  $a$ ,  $b$  via an optical circulator, the Michelson interferometer is formally equivalent to the Mach-Zehnder interferometer. The input output relations are identical as in Eq. (27) with  $\varphi = 2(\varphi_b - \varphi_a)$ , as the light acquires the relative phase twice—traveling both to and from the end mirrors.

Consider now the Fabry-Pérot interferometer depicted in Fig. 3b. We assume for simplicity that both mirrors have the same power transmission coefficient  $T$  and the phase  $\theta$  is acquired while the light travels from one mirror to the other inside the interferometer. The

resulting input-output relation reads:

$$\begin{pmatrix} \hat{a}' \\ \hat{b}' \end{pmatrix} = \frac{1}{(2-T)i\sin\theta - T\cos\theta} \times \begin{pmatrix} 2i\sqrt{1-T}\sin\theta & T \\ T & 2i\sqrt{1-T}\sin\theta \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \quad (45)$$

and up to an irrelevant global phase may be rewritten as:

$$\begin{pmatrix} \hat{a}' \\ \hat{b}' \end{pmatrix} = \begin{pmatrix} \cos\varphi/2 & i\sin\varphi/2 \\ i\sin\varphi/2 & \cos\varphi/2 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \quad (46)$$

$$\varphi = 2 \arcsin \left( \frac{T}{\sqrt{T^2 + 4(1-T)\sin^2\theta}} \right).$$

In terms of the angular momentum operators the above transformation is simply a  $\varphi$ -rotation around the  $x$  axis. Thus, up to a change of the rotation axis, the action of the Fabry-Pérot interferometer with phase delay  $\theta$  is equivalent to the one of the MZ interferometer with phase delay  $\varphi$ . Knowing the formulas for the estimation precision of  $\varphi$  in the MZ interferometer, we can easily calculate the corresponding estimation precision of  $\theta$  via the error propagation formula obtaining:

$$\Delta\theta = \frac{\Delta\varphi}{|\frac{\partial\varphi}{\partial\theta}|} = \Delta\varphi \frac{T^2 + 4(1-T)\sin^2\theta}{4T\sqrt{1-T}\cos\theta}. \quad (47)$$

As a consequence, the above expression allows to translate all the results derived for the MZ interferometer to the Fabry-Pérot case.

Ramsey interferometry is a popular technique for performing precise spectroscopic measurements of atoms. It is widely used in atomic clock setups, where it allows to lock the frequency of an external source of radiation,  $\omega$ , to a selected atomic transition frequency,  $\omega_0$ , between the single-atom excited and ground states,  $|e\rangle$  and  $|g\rangle$  (Diddams *et al.*, 2004). In a typical Ramsey interferometric experiment  $N$  atoms are initially prepared in the ground state and are subsequently subjected to a  $\pi/2$  Rabi pulse, which transforms each of them into an equally weighted superposition of ground and excited states. Afterwards, they evolve freely for time  $t$ , before finally being subjected to a second  $\pi/2$  pulse, which in the ideal case of  $\omega_0 = \omega$  would put all the atoms in the excited state. In case of any frequency mismatch, the probability for an atom to be measured in an excited state is  $\cos^2(\varphi/2)$ , where  $\varphi = (\omega_0 - \omega)t$ . Hence, by measuring the number of atoms in the excited state, one can estimate  $\varphi$  and consequently knowing  $t$  the frequency difference  $\omega_0 - \omega$ . Treating two-level atoms as spin-1/2 particles with their two levels corresponding to up and down projections of the spin  $z$  component, we may introduce the total spin operators  $\hat{J}_i = \frac{1}{2} \sum_{k=0}^N \hat{\sigma}_i^{(k)}$ ,  $i = x, y, z$ , where  $\hat{\sigma}_i^{(k)}$  are standard Pauli sigma matrices acting on the  $k$ -th particle. Evolution of a general input state can be written as:

$$|\psi_\varphi\rangle = e^{-i\hat{J}_x\pi/2} e^{i\hat{J}_z\varphi} e^{-i\hat{J}_x\pi/2} |\psi\rangle_{\text{in}}. \quad (48)$$



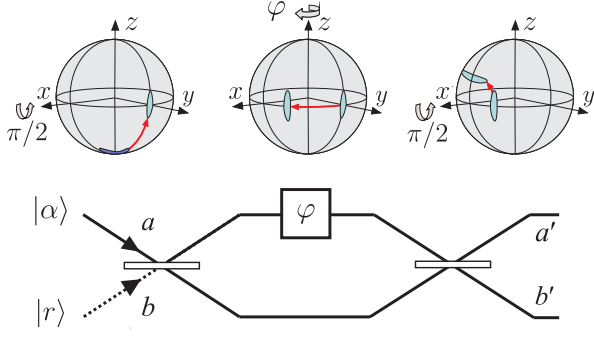


FIG. 4 Formal equivalence of Ramsey and MZ interferometry. The two atomic levels which are used in the Ramsey interferometry play analogous roles as the two arms of the interferometer, while the  $\pi/2$  pulses equivalent from a mathematical point of view to the action of the beam-splitters. Quantum-enhancement in Ramsey interferometry may be obtained by preparing the atoms in a spin-squeezed input state reducing the variances of the relevant total angular momentum operators, in a similar fashion as using squeezed states of light leads to an improved sensitivity in the MZ interferometry.

which is completely analogous to the MZ transformation (30) with  $\pi/2$  pulses playing the role of the beam-splitters, see Fig. 4. The total spin  $z$  operator can be written as  $\hat{J}_z = 2(\hat{n}_g - \hat{n}_e)$ , where  $\hat{n}_g$ ,  $\hat{n}_e$  denote the ground and excited state atom number operators respectively. Therefore, measurement of  $\hat{J}_z$  is equivalent to the measurement of the difference of excited and non-excited atoms analogously to the optical case where it corresponded to the measurement of photon-number difference at the two output ports of the interferometer. Fluctuation of the number of atoms measured limits the estimation precision and in case of uncorrelated atoms is referred to as the *projection noise*, which may be regarded as an analog of the optical shot noise. An important difference from the optical case, though, is that when dealing with atoms we are restricted to consider states of definite particle-number. Thus, there is no exact analogue of coherent or squeezed states that we consider in the photonic case. We can therefore regard atomic Ramsey interferometry as a special case of the MZ interferometry with inputs restricted to states of definite particle-number, discussed in Sec. III.E, and further relate the precision of estimating the frequency difference to the precision of phase estimation via

$$\Delta\omega = \frac{\Delta\varphi}{t}. \quad (49)$$

Beating the projection noise requires the input state of the atoms to share some particle entanglement. From an experimental point of view the most promising class of states are the so-called *one-axis or two-axis spin-squeezed states*, which may be realized in BEC and atomic systems interacting with light (Kitagawa and Ueda, 1993; Ma *et al.*, 2011). In fact, these states may be regarded as a definite particle-number analogues of optical squeezed

states. In particular, starting with atoms in a ground state the two-axis spin-squeezed states may be obtained via:

$$|\psi_\chi\rangle = e^{-\frac{\chi}{2}(\hat{J}_+^2 - \hat{J}_-^2)}|g\rangle^{\otimes N}, \quad (50)$$

where  $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$ . The above formula resembles the definition of an optical squeezed state given in Eq. (10), where  $\hat{a}$ ,  $\hat{a}^\dagger$  operators are replaced by  $\hat{J}_-$ ,  $\hat{J}_+$ . With an appropriate choice of squeezing strength  $\chi$  as a function of  $N$  it is possible to achieve the Heisenberg scaling of precision  $\Delta\omega \propto 1/N$  (Ma *et al.*, 2011; Wineland *et al.*, 1994).

Analogous schemes may also be implemented in BEC (Cronin *et al.*, 2009; Gross *et al.*, 2010). In particular, BEC opens a way of realizing a specially appealing matter-wave interferometry, in which, similarly to the optical interferometers, the matter-wave is split into two spatial modes, evolves and finally interferes resulting in a spatial fringe pattern that may be used to estimate the relative phase acquired by the atoms (Shin *et al.*, 2004). Such a scheme may potentially find applications in precise measurements of the gravitational field (Anderson and Kasevich, 1998). Yet, in this case, the detection involves measurements of positions of the atoms forming the interference fringes, what makes the estimation procedure more involved than in the simple MZ scheme (Chwedeńczuk *et al.*, 2012, 2011), but eventually the precisions for the optimal estimation schemes should coincide with the ones obtained for the MZ interferometry.

Finally, we should also mention that atomic ensembles interacting with light are excellent candidates for ultra-precise magnetometers (Budker and Romalis, 2007). Collective magnetic moment of atoms rotates in the presence of magnetic field to be measured, what again can be seen as an analogue of the MZ transformation on  $\hat{J}_i$  in Eq. (30). The angle of the atomic magnetic-moment rotation is determined by sending polarized light which due to the Faraday effect is rotated proportionally to the atomic magnetic-moment component in the direction of the light propagation. In standard scenarios, the ultimate precision will be affected by both the atomic projection noise, due to characteristic uncertainties of the collective magnetic-moment operator for uncorrelated atoms, and the light shot noise. The quantum enhancement of precision may again be achieved by squeezing the atomic states (Sewell *et al.*, 2012; Wasilewski *et al.*, 2010) as well as by using the non-classical states of light (Horrom *et al.*, 2012), what in both cases allows to go beyond the projection and shot-noise limits.

#### IV. ESTIMATION THEORY

In this section we review the basics of both classical and quantum estimation theory. We present Fisher Information and Bayesian approaches to determining the

optimal estimation strategies and discuss tools particularly useful for analysis of optical interferometric setups.

### A. Classical parameter estimation

The essential question that has been addressed by statisticians long before the invention of quantum mechanics is how to most efficiently extract information from a given data set, which is determined by some non-deterministic process (Kay, 1993; Lehmann and Casella, 1998).

In a typical scenario we are given an  $N$ -point data set  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$  which is a realization of  $N$  independent identically distributed random variables,  $X^N$ , each distributed according to a common Probability Density Function (PDF),  $p_\varphi(X)$ , that depends on an unknown parameter  $\varphi$  we wish to determine. Our goal is to construct an *estimator*  $\tilde{\varphi}_N(\mathbf{x})$  that should be interpreted as a function which outputs the most accurate estimate of the parameter  $\varphi$  based on a given data set. Importantly, as the estimator  $\tilde{\varphi}_N$  is build on a sample of random data, it is a random variable itself and the smaller are its fluctuations around the true value  $\varphi$  the better it is.

Typically, two approaches to the problem of the choice of the optimal estimator are undertaken. In the so-called frequentist or classical approach,  $\varphi$  is assumed to be a deterministic variable with an unknown value that, if known, could in principle be stated to any precision. In this case, one of the basic tools in studying optimal estimation strategies is the Fisher information, and hence we will refer to this approach as the *Fisher information* approach. In contrast, when following the *Bayesian* paradigm, the estimated parameter is a random variable itself that introduces some intrinsic error that accounts for the lack of knowledge about  $\varphi$  we possess *prior* to performing the estimation. We describe both approaches in detail below.

#### 1. Fisher Information approach

In this approach  $p_\varphi(X)$  is regarded as a family of PDFs parametrized by  $\varphi$ —the parameter to be estimated based on the registered data  $\mathbf{x}$ . The performance of a given estimator  $\tilde{\varphi}_N(\mathbf{x})$  is quantified by the Mean Square Error (MSE) deviation from the true value  $\varphi$ :

$$\Delta^2 \tilde{\varphi}_N = \left\langle (\tilde{\varphi}_N(\mathbf{x}) - \varphi)^2 \right\rangle = \int d^N \mathbf{x} p_\varphi(\mathbf{x}) (\tilde{\varphi}_N(\mathbf{x}) - \varphi)^2. \quad (51)$$

A desired property for an estimator is that it is *unbiased*:

$$\langle \tilde{\varphi}_N \rangle = \int d^N \mathbf{x} p_\varphi(\mathbf{x}) \tilde{\varphi}_N(\mathbf{x}) = \varphi, \quad (52)$$

so that on average it yields the true parameter value. The optimal unbiased estimator is the one that minimizes

$\Delta^2 \tilde{\varphi}_N$  for all  $\varphi$ . Looking for the optimal estimator may be difficult and it may even be the case that there is no single estimator that minimizes the MSE for all  $\varphi$ .

Still, one may always construct the so-called *Cramer-Rao Bound* (CRB) that lower-bounds the MSE of any unbiased estimator  $\tilde{\varphi}_N$  (see e.g. (Kay, 1993) for a review):

$$\Delta^2 \tilde{\varphi}_N \geq \frac{1}{N F[p_\varphi]}, \quad (53)$$

where  $F$  is the *Fisher Information* (FI), and can be expressed using one of the formulas below:

$$F[p_\varphi] = \int dx \frac{1}{p_\varphi(x)} \left[ \frac{\partial p_\varphi(x)}{\partial \varphi} \right]^2 = \left\langle \left( \frac{\partial}{\partial \varphi} \ln p_\varphi \right)^2 \right\rangle = - \left\langle \frac{\partial^2}{\partial \varphi^2} \ln p_\varphi \right\rangle, \quad (54)$$

The basic intuition is that the bigger the FI is the higher estimation precision may be expected. The FI is non-negative and additive for uncorrelated events, so that  $F[p_\varphi^{(1,2)}] = F[p_\varphi^{(1)}] + F[p_\varphi^{(2)}]$ , for  $p_\varphi^{(1,2)}(x_1, x_2) = p_\varphi^{(1)}(x_1) p_\varphi^{(2)}(x_2)$  and in particular:  $F[p_\varphi^N] = N F[p_\varphi]$ , which can be easily verified using the last expression in definition (54). The FI is straightforward to calculate and once an estimator is found that saturates the CRB it is guaranteed to be optimal. In general estimators saturating the CRB are called *efficient*. The sufficient and necessary condition for efficiency is the following condition on the PDF and the estimator (Kay, 1993):

$$\frac{\partial}{\partial \varphi} \ln p_\varphi(\mathbf{x}) = N F[p_\varphi] (\tilde{\varphi}_k(\mathbf{x}) - \varphi). \quad (55)$$

An estimator  $\tilde{\varphi}$  satisfying the above equality exist only for a special class of PDFs belonging to the so called *exponential family* of PDFs, for which:

$$\ln p_\varphi(x) = a(\varphi) + b(x) + c(\varphi)d(x), \quad \frac{a'(\varphi)}{c'(\varphi)} = -\varphi, \quad (56)$$

where  $a(\varphi)$ ,  $c(\varphi)$  and  $b(x)$ ,  $d(x)$  are arbitrary functions and primes denote differentiation over  $\varphi$ . In general, however, the saturability condition cannot be met.

Note that in general FI is a function of  $\varphi$ , so that depending on the true value of the parameter, the CRB puts weaker or stronger constraints on the minimal MSE. Actually, one is not always interested in the optimal estimation strategy that is valid *globally*—for any potential value of  $\varphi$ —but may want to design a protocol that works optimally for  $\varphi$  confined to some small parameter range. In this case one can take a *local approach* and analyze the CRB at a given point  $\varphi = \varphi_0$ . Formally derivation of the CRB at a given point requires only a weaker *local unbiasedness* condition:

$$\left. \frac{\partial}{\partial \varphi} \langle \tilde{\varphi}_N \rangle \right|_{\varphi=\varphi_0} = 1 \quad (57)$$

at a given parameter value  $\varphi_0$ . FI at  $\varphi_0$  is a *local* quantity that depends only on  $p_\varphi(X)|_{\varphi=\varphi_0}$  and  $\left.\frac{\partial p_\varphi(X)}{\partial \varphi}\right|_{\varphi=\varphi_0}$ , as explicitly stated in Eq. (54). As a result, the FI is sensitive to changes of the PDF of the first order in  $\delta\varphi = \varphi - \varphi_0$ . Looking for the optimal locally unbiased estimator at a given point  $\varphi_0$  makes sense provided one has a substantial prior knowledge that the true value of  $\varphi$  is close to  $\varphi_0$ . This may be the case if the data is obtained from a well controlled physical system subjected to small external fluctuations or if some part of the data had been used for preliminary estimation narrowing the range of compatible  $\varphi$  to a small region around  $\varphi_0$ . In this case, even if the condition for saturability of the CRB cannot be met, it still may be possible to find a locally unbiased estimator which will saturate the CRB at least at a given point  $\varphi_0$ . The explicit form of the estimator can easily be derived from Eq. (55) by substituting  $\varphi = \varphi_0$ :

$$\tilde{\varphi}_N^{\varphi_0}(\mathbf{x}) = \varphi_0 + \frac{1}{NF[p_\varphi]} \left. \frac{\partial \ln p_\varphi(\mathbf{x})}{\partial \varphi} \right|_{\varphi=\varphi_0}. \quad (58)$$

Fortunately, difficulties in saturating the CRB are only present in the finite- $N$  regime. In the asymptotic limit of infinitely many repetitions of an experiment, or equivalently, for an infinitely large sample,  $N \rightarrow \infty$ , a particular estimator called the *Maximum Likelihood* (ML) estimator saturates the CRB (Kay, 1993; Lehmann and Casella, 1998). The ML estimator formally defined as

$$\tilde{\varphi}_N^{\text{ML}}(\mathbf{x}) = \underset{\varphi}{\operatorname{argmax}} p_\varphi(\mathbf{x}) \quad (59)$$

is a function that for a given instance of outcomes,  $\mathbf{x}$ , outputs the value of parameter for which this data sample is the most probable. For finite  $N$  the ML estimator is in general biased, but becomes unbiased asymptotically  $\lim_{N \rightarrow \infty} \langle \varphi_N^{\text{ML}}(\mathbf{x}) \rangle = \varphi$  and saturates the CR bound  $\lim_{N \rightarrow \infty} N \Delta^2 \tilde{\varphi}_N^{\text{ML}}(\mathbf{x}) = F[p_\varphi]$ .

## 2. Bayesian approach

In this approach, the parameter to be estimated,  $\varphi$ , is assumed to be a random variable that is distributed according to a *prior* PDF,  $p(\varphi)$ , representing the knowledge about  $\varphi$  one possesses before performing the estimation, while  $p(x|\varphi)$  denotes the conditional probability of obtaining result  $x$  for parameter value  $\varphi$ . Notice, a subtle change in notation from  $p_\varphi(x)$  in the FI approach to  $p(x|\varphi)$  in the Bayesian approach reflecting the change in the role of  $\varphi$  which is a parameter in the FI approach and a random variable in the Bayesian approach. If we stick to the MSE as a cost function, we say that the estimator  $\tilde{\varphi}_N(\mathbf{x})$  is optimal if it minimizes the average MSE

$$\langle \Delta^2 \tilde{\varphi}_N \rangle = \iint d\varphi d\mathbf{x} p(\mathbf{x}|\varphi) p(\varphi) (\tilde{\varphi}_N(\mathbf{x}) - \varphi)^2, \quad (60)$$

which, in contrast to Eq. (51), is also averaged over all the values of the parameter with the Bayesian prior  $p(\varphi)$ . Making use of the Bayes theorem we can rewrite the above expression in the form

$$\langle \Delta^2 \tilde{\varphi}_N \rangle = \int d\mathbf{x} p(\mathbf{x}) \left[ \int d\varphi p(\varphi|\mathbf{x}) (\tilde{\varphi}_N(\mathbf{x}) - \varphi)^2 \right]. \quad (61)$$

From the above formula it is clear that the optimal estimator is the one that minimizes terms in square bracket for each  $\mathbf{x}$ . Hence, we can explicitly derive the form of the *Minimum Mean Squared Error* (MMSE) estimator

$$\begin{aligned} \frac{\partial}{\partial \tilde{\varphi}_N} \int d\varphi p(\varphi|\mathbf{x}) (\tilde{\varphi}_N(\mathbf{x}) - \varphi)^2 = 0 \implies \\ \tilde{\varphi}_N^{\text{MMSE}}(\mathbf{x}) = \int d\varphi p(\varphi|\mathbf{x}) \varphi = \langle \varphi \rangle_{p(\varphi|\mathbf{x})} \end{aligned} \quad (62)$$

which simply corresponds to the *average* value of the parameter computed with respect to the *posterior* PDF,  $p(\varphi|\mathbf{x})$ . The posterior PDF represents the knowledge we possess about the parameter after inferring the information about it from the sampled data  $\mathbf{x}$ . The corresponding MMSE reads:

$$\begin{aligned} \langle \Delta^2 \tilde{\varphi}_N \rangle &= \int d\mathbf{x} p(\mathbf{x}) \left[ \int d\varphi p(\varphi|\mathbf{x}) \left( \varphi - \langle \varphi \rangle_{p(\varphi|\mathbf{x})} \right)^2 \right] \\ &= \int d\mathbf{x} p(\mathbf{x}) \Delta^2 \varphi|_{p(\varphi|\mathbf{x})}, \end{aligned} \quad (63)$$

so that it may be interpreted as the variance of the parameter  $\varphi$  computed again with respect to the posterior PDF  $p(\varphi|\mathbf{x})$  that is averaged over all the possible outcomes.

The optimal estimation strategy within the Bayesian approach depends explicitly on the prior PDF assumed. If either the prior PDF will be very sensitive to variations of  $\varphi$  or the physical model will predict the data to be weakly affected by any parameter changes, so that  $p(\mathbf{x}|\varphi)p(\varphi) \approx p(\varphi)$ , the minimal  $\Delta^2 \tilde{\varphi}_N$  will be predominantly determined by the prior distribution  $p(\varphi)$  and the sampled data will have limited effect on the estimation process. Therefore, it is really important in the Bayesian approach to choose an appropriate prior PDF that, on one hand, should adequately represent our knowledge about the parameter before the estimation, but, on the other, its choice should not dominate the information obtained from the data.

In principle, nothing prevents us to consider more general *cost functions*,  $C(\tilde{\varphi}, \varphi)$ , that in some situations may be more suitable than the squared error. The corresponding optimal estimator will be the one that minimizes the average cost function

$$\langle C \rangle = \iint d\varphi d\mathbf{x} p(\varphi) p(\mathbf{x}|\varphi) C(\tilde{\varphi}_N(\mathbf{x}), \varphi). \quad (64)$$

In the context of optical interferometry the estimated quantity of interest  $\varphi$  is the phase which is a circular

parameter, i.e. may be identified with a point on a circle or more formally as an element of the circle group  $U(1)$  and in particular  $\varphi \equiv \varphi + 2n\pi$ . Following (Holevo, 1982) the cost function should respect the parameter *topology* and the squared error is clearly not the proper choice. We require the cost function to be symmetric,  $C(\tilde{\varphi}, \varphi) = C(\varphi, \tilde{\varphi})$ , *group invariant*, i.e.  $\forall \phi \in U(1): C(\tilde{\varphi} + \phi, \varphi + \phi) = C(\tilde{\varphi}, \varphi)$ , and *periodic*,  $C(\varphi + 2n\pi, \tilde{\varphi}) = C(\varphi, \tilde{\varphi})$ . This restricts the class of cost functions to:

$$C(\tilde{\varphi}, \varphi) = C(\delta\varphi) = \sum_{n=0}^{\infty} c_n \cos[n\delta\varphi] \quad \text{with } \delta\varphi = \tilde{\varphi} - \varphi. \quad (65)$$

Furthermore, we require  $C(\delta\varphi)$  must rise monotonically from  $C(0)=0$  at  $\delta\varphi=0$  to some  $C(\pi)=C_{\max}$  at  $\delta\varphi=\pi$ , so that  $C'(\delta\varphi) \geq 0$ , so that the coefficients  $c_n$  must fulfill following constraints:

$$\begin{aligned} \sum_{n=0}^{\infty} c_n &= 0, & \sum_{n=0}^{\infty} (-1)^n c_n &= C_{\max}, \\ \sum_{n=1}^{\infty} n^2 c_n &\leq 0, & \sum_{n=1}^{\infty} n^2 (-1)^n c_n &\geq 0, \end{aligned} \quad (66)$$

which may be satisfied by imposing  $\forall_{n>0}: c_n \leq 0$  and taking  $c_n$  to decay at least quadratically with  $n$ . Lastly, for the sake of compatibility we would like the cost function to approach the standard variance for small  $\delta\varphi$  so that  $C(\delta\varphi) = \delta\varphi^2 + O(\delta\varphi^4)$ , which is equivalent to  $\sum_{n=1}^{\infty} n^2 c_n = -2$ .

In all the Bayesian estimation problems considered in this work, we will consider the simplest cost function that satisfies all above-mentioned conditions with  $c_0 = -c_1 = 2$ ,  $\forall_{n>1}: c_n = 0$  which reads explicitly:

$$C(\tilde{\varphi}, \varphi) = 4 \sin^2\left(\frac{\tilde{\varphi} - \varphi}{2}\right). \quad (67)$$

Following the same argumentation as described in Eq. (62) when minimizing the averaged MSE, one may prove that for the above chosen cost function the average cost  $\langle C \rangle$  is minimized if an estimator,  $\tilde{\varphi}_N^C(\mathbf{x})$ , can be found that for any possible data sample  $\mathbf{x}$  collected satisfies the condition

$$\int d\varphi \, p(\varphi|\mathbf{x}) \sin(\tilde{\varphi}_N^C(\mathbf{x}) - \varphi) = 0. \quad (68)$$

### 3. Example: Transmission coefficient estimation

In order to illustrate the introduced concepts let us consider a simple model of parameter estimation, where a single photon impinges on the beamsplitter with power transmission and reflectivity equal respectively  $\mathbf{p}$  and  $\mathbf{q}=1-\mathbf{p}$ . The experiment is repeated  $N$  times and based on the data obtained—number of photons transmitted and the number of photons reflected—the goal is to estimate the transmission coefficient  $\mathbf{p}$ . This problem is

equivalent to a coin-tossing experiment, where we assume an unfair coin, which “heads” and “tails” occurring with probabilities  $\mathbf{p}$  and  $\mathbf{q}=1-\mathbf{p}$  respectively.

Probability that  $n$  out of  $N$  photons get transmitted is governed by the binomial distribution

$$p_{\mathbf{p}}^N(n) = \binom{N}{n} \mathbf{p}^n (1-\mathbf{p})^{N-n}. \quad (69)$$

The FI equals  $F[p_{\mathbf{p}}^N] = N/[\mathbf{p}(1-\mathbf{p})]$  and hence the CRB imposes a lower bound on the achievable estimation variance:

$$\Delta^2 \tilde{\mathbf{p}}_N \geq \frac{\mathbf{p}(1-\mathbf{p})}{N}. \quad (70)$$

Luckily the binomial probability distribution belongs to the exponential family of PDFs specified in Eq. (56), and by inspecting saturability condition Eq. (55) it may be easily checked, that the simple estimator  $\tilde{\mathbf{p}}_N(n) = n/N$  saturates the CRB. It is also worth mentioning that the optimal estimator also coincides with the ML estimator, hence in this case the ML estimator is optimal also for finite  $N$  and not only in the asymptotic regime.

In the context of optical interferometry, we will deal with an analogous situation, where photons are sent into one input port of an interferometer and  $\mathbf{p}, \mathbf{q}$  correspond to the probabilities of detecting a photon in one of the two output ports. For a Mach-Zehnder interferometer, see Sec. III, the probabilities depend on the relative phase delay difference between the arms of an interferometer  $\varphi$ :  $\mathbf{p} = \sin^2(\varphi/2)$ ,  $\mathbf{q} = \cos^2(\varphi/2)$ , which is the actual parameter of interest. Probability distribution as a function of  $\varphi$  then reads:

$$p_{\varphi}^N(n) = \binom{N}{n} \sin(\varphi/2)^{2n} \cos(\varphi/2)^{2(N-n)} \quad (71)$$

and the corresponding FI and the CRB take the form

$$F[p_{\varphi}^N] = N, \quad \Delta^2 \tilde{\varphi}_N \geq \frac{1}{N}. \quad (72)$$

Interestingly, FI does not depend on the actual value  $\varphi$ , what suggests that the achievable estimation precision may be independent of the actual parameter value. However, for such a parametrization, the CRB saturability condition Eq. (55) does not hold and there is no unbiased estimator saturating the bound. Nevertheless, using Eq. (58), we may still write a locally unbiased estimator saturating CRB at  $\varphi_0$ :  $\tilde{\varphi}_N(n) = \varphi_0 - \tan(\varphi_0/2) + 2n/(N \sin \varphi_0)$ , which is possible provided  $\sin \varphi_0 \neq 0$ .

Since the CRB given in (72) can only be saturated locally it is worth looking at the MLE which we know will perform optimally in the asymptotic regime  $N \rightarrow \infty$ . Solving Eq. (59) we obtain

$$\tilde{\varphi}_N(n) = \operatorname{argmax}_{\varphi} \ln p_{\varphi}^N(n) = \pm 2 \arctan \sqrt{\frac{n}{N-n}} \quad (73)$$



and in general we notice that there are two equivalent maxima. This ambiguity is simply the result of invariance of the PDF  $p_\varphi(n)$  with respect to the change  $\varphi \rightarrow -\varphi$  and might have been expected. Hence in practice we would need some additional information, possibly coming from prior knowledge or other observations, in order to distinguish between this two cases and be entitled to claim that the MLE saturates the CRB asymptotically for all  $\varphi$ .

We now analyze the same estimation problem employing the Bayesian approach. Let us first consider the case where  $\mathbf{p}$  is the parameter to be estimated and the relevant conditional probability  $p^N(n|\mathbf{p}) = p_{\mathbf{p}}^N(n)$  is given by Eq. (69). Choosing a flat prior distribution  $p(\mathbf{p}) = 1$  and the mean square as a cost function, we find using Eq. (62) and Eq. (63) that the MMSE estimator and the corresponding minimal averaged MSE equal:

$$\tilde{\mathbf{p}}_N(n) = \frac{n+1}{N+2}, \quad \langle \Delta^2 \tilde{\mathbf{p}}_N \rangle = \frac{1}{6(N+2)}, \quad (74)$$

which may be compared with the previously discussed FI approach where the optimal estimator was  $\tilde{\mathbf{p}}(n) = n/N$  and the resulting variance when averaged over all  $\mathbf{p}$  would yield  $\langle \Delta^2 \tilde{\mathbf{p}}_N \rangle = \int_0^1 dp \frac{p(p-1)}{N} = 1/(6N)$ . Hence, in the limit  $N \rightarrow \infty$  these results converge to the ones obtained previously in the FI approach. This is a typical situation that in the case of large amount of data the two approaches yield equivalent results (van der Vaart, 2000).

We now switch to  $\varphi$  parametrization and consider  $p^N(n|\varphi) = p_\varphi^N(n)$  as in Eq. (71). Assuming flat prior distribution  $p(\varphi) = 1/(2\pi)$  and the previously introduced natural cost function in the case of circular parameter  $C(\tilde{\varphi}, \varphi) = 4 \sin^2[(\tilde{\varphi} - \varphi)/2]$ , due to  $\pm\varphi$  estimation ambiguity we realize that the condition for the optimal estimator given in Eq. (68) is satisfied for a trivial estimator  $\tilde{\varphi}^C(n) = 0$  which does not take into account the measurement results at all. This can be understood once we realize that the ambiguity in the sign of estimated phase  $\varphi$  and the possibility of estimating the phase with the wrong sign is worse than not taking into account the measured data at all. In order to obtain a more interesting result we need to consider a subset of possible values of  $\varphi$  over which the reconstruction is not ambiguous. If we choose  $\varphi \in [0, \pi)$ , and the corresponding prior  $p(\varphi) = 1/\pi$ , Eq. (68) yields the optimal form of the estimator and the corresponding minimal cost calculated according to Eq. (64) are as follows:

$$\tilde{\varphi}_N^C(n) = \arctan\left(\frac{2}{f(N,n)}\right) \quad (75)$$

$$\langle C \rangle = 2 \left( 1 - \frac{1}{\pi(N+1)} \sum_{n=0}^N \sqrt{4 + f(N,n)^2} \right), \quad (76)$$

where  $f(N,n) = (N-2n)(n-1/2)!(N-n-1/2)!/n!(N-n)!$ . Despite its complicated form the above formu-

las simplify in the limit  $N \rightarrow \infty$ . The optimal estimator approaches the ML estimator  $\tilde{\varphi}_N^C(n) \approx 2 \times \arctan \sqrt{n/(N-n)}$  while the average cost function approaches  $\langle C \rangle \approx 1/N$  indicating saturation of the CRB and confirming again that in the regime of many experimental repetitions the two discussed approaches coincide.

## B. Quantum parameter estimation

In a quantum estimation scenario the parameter  $\varphi$  is encoded in a *quantum state*  $\rho_\varphi$  which is subject to a quantum measurement  $M_{\mathbf{x}}$  yielding measurement result  $\mathbf{x}$  with probability  $p_\varphi(\mathbf{x}) = \text{Tr}\{\rho_\varphi M_{\mathbf{x}}\}$ . The estimation strategy is complete once an estimator function  $\tilde{\varphi}(\mathbf{x})$  is given ascribing estimated parameters to particular measurement results. The quantum measurement may be a standard projective von-Neumann measurement,  $M_{\mathbf{x}} M_{\mathbf{x}'} = M_{\mathbf{x}} \delta_{\mathbf{x},\mathbf{x}'}$ , or a *generalized measurement* where the measurement operators form a *Positive Operator Valued Measure* (POVM) with the only constraints being  $M_{\mathbf{x}} \geq 0$ ,  $\int d\mathbf{x} M_{\mathbf{x}} = \mathbb{1}$  (Bengtsson and Zyczkowski, 2006; Nielsen and Chuang, 2000). Establishing the optimal estimation strategy corresponds then not only to the most accurate inference of the parameter value from the data, but also to a non-trivial optimization over the class of all POVMs to find the measurement scheme maximizing the precision. However, as soon as we decide on a particular measurement scheme, we obtain a model that determines probabilities describing the sampled data  $p_\varphi(\mathbf{x})$  and their quantum mechanical origin becomes irrelevant. Hence, the estimation problem becomes then fully classical and all the techniques developed in Sec. IV.A apply.

### 1. Quantum Fisher Information approach

The problem of determining the optimal measurement scheme for a particular estimation scenario is non-trivial. Fortunately analogously as in the classical estimation it is relatively easy to obtain useful lower bounds on the minimal MSE. The *Quantum Cramér-Rao Bound* (QCRB) (Braunstein and Caves, 1994; Helstrom, 1976; Holevo, 1982) is a generalization of the classical CRBs (53), which lower bounds the variance of estimation for all possible locally unbiased estimators and the most general POVM measurements<sup>1</sup>:

$$\Delta^2 \varphi_N \geq \frac{1}{N F_Q[\rho_\varphi]} \quad \text{with} \quad F_Q[\rho_\varphi] = \text{Tr}\left\{\rho_\varphi L[\rho_\varphi]^2\right\} \quad (77)$$

<sup>1</sup> For clarity of notation, in what follows we drop the tilde symbol when writing estimator variance.

where  $F_Q$  is the *Quantum Fisher Information* (QFI) while the Hermitian operator  $L[\rho_\varphi]$  is the so called *Symmetric Logarithmic Derivative* (SLD), which can be unambiguously defined for any state  $\rho_\varphi$  via the relation  $\dot{\rho}_\varphi = \frac{1}{2}(\rho_\varphi L[\rho_\varphi] + L[\rho_\varphi] \rho_\varphi)$ . Crucially, QFI is solely determined by the dependence of  $\rho_\varphi$  on the estimated parameter, and hence allows to analyze parameter sensitivity of given probe states without considering any particular measurements nor estimators. Explicitly, the SLD when written in the eigenbasis of  $\rho_\varphi = \sum_i \lambda_i(\varphi) |e_i(\varphi)\rangle\langle e_i(\varphi)|$  reads:

$$L[\rho_\varphi] = \sum_{i,j} \frac{2 \langle e_i(\varphi) | \dot{\rho}_\varphi | e_j(\varphi) \rangle}{\lambda_i(\varphi) + \lambda_j(\varphi)} |e_i(\varphi)\rangle\langle e_j(\varphi)|, \quad (78)$$

where the sum is taken over the terms with non-vanishing denominator. Analogously to the FI (54), the QFI is an additive quantity when calculated on product states and in particular  $F_Q[\rho_\varphi^{\otimes N}] = N F_Q[\rho_\varphi]$ . Thus the  $N$  term in the denominator of Eq. (77) may be equivalently interpreted as the number of independent repetitions of an experiment with a state  $\rho_\varphi$  to form the data sample  $\mathbf{x}$  of size  $N$ , or a single shot experiment with a multi-party state  $\rho_\varphi^{\otimes N}$ .

Crucially, as proven in (Braunstein and Caves, 1994; Nagaoka, 2005), there always exist a measurement strategy—a projection measurement in the eigenbasis of the SLD—for which the FI calculated for the resulting probability distribution equals the QFI, and consequently the bounds (53) and (77) coincide. Hence the issue of saturability of the QCRB amounts to the problem of saturability of the corresponding classical CRB. As discussed in detail in Sec. IV.A.1, the bound is therefore globally saturable for a special class of probability distributions belonging to the so called exponential family, and if this is not the case the saturability is achievable either in the asymptotic limit of many independent experiments  $N \rightarrow \infty$  or in the local approach when one estimates small fluctuation of the parameter in the vicinity of a known value  $\varphi_0$ .

For pure states,  $\rho_\varphi = |\psi_\varphi\rangle\langle\psi_\varphi|$ , the QFI in Eq. (77) simplifies to

$$F_Q[|\psi_\varphi\rangle] = 4 \left( \left\langle \dot{\psi}_\varphi | \dot{\psi}_\varphi \right\rangle - \left| \left\langle \dot{\psi}_\varphi | \psi_\varphi \right\rangle \right|^2 \right), \quad |\dot{\psi}_\varphi\rangle = \frac{d|\psi_\varphi\rangle}{d\varphi}. \quad (79)$$

Yet, for general mixed states, calculation of the QFI involves diagonalization of the quantum state  $\rho_\varphi$ , in order to calculate the SLD, and becomes tedious for probe states living in highly dimensional Hilbert spaces.

Interestingly, the QFI may be alternatively calculated by considering purifications  $|\Psi_\varphi\rangle$  of a given family of mixed states on an extended Hilbert space  $\rho_\varphi = \text{Tr}_E\{|\Psi_\varphi\rangle\langle\Psi_\varphi|\}$ , where by  $E$  we denote an ancillary space needed for the purification. It has been proven by Escher *et al.* (2011) that the QFI of any  $\rho_\varphi$  is equal to the small-

est QFI of its purifications  $|\Psi_\varphi\rangle$

$$F_Q[\rho_\varphi] = \min_{|\Psi_\varphi\rangle} F_Q[|\Psi_\varphi\rangle] = 4 \min_{|\Psi_\varphi\rangle} \left\{ \left\langle \dot{\Psi}_\varphi | \dot{\Psi}_\varphi \right\rangle - \left| \left\langle \dot{\Psi}_\varphi | \Psi_\varphi \right\rangle \right|^2 \right\}. \quad (80)$$

Even though minimization over all purification may still be challenging, the above formulation may easily be employed in deriving upper bounds on QFI by considering some class of purifications. Since upper bounds on QFI translate to lower bounds on estimation uncertainty this approach proved useful in deriving bounds in quantum metrology in presence of decoherence (Escher *et al.*, 2012, 2011).

Independently, in (Fujiwara and Imai, 2008) another purification-based QFI definition has been constructed:

$$F_Q[\rho_\varphi] = 4 \min_{|\Psi_\varphi\rangle} \left\langle \dot{\Psi}_\varphi | \dot{\Psi}_\varphi \right\rangle. \quad (81)$$

Despite apparent difference, Eqs. (80) and (81) are equivalent and one can prove that any purification minimizing one of them is likewise optimal for the other causing the second term of Eq. (80) to vanish. Although for any suboptimal  $|\Psi_\varphi\rangle$  Eq. (80) must provide a strictly tighter bound on QFI than Eq. (81), the latter definition, owing to its elegant form allows for a direct and efficient procedure for derivation of the precision bounds in quantum metrology (Demkowicz-Dobrzański *et al.*, 2012; Kołodyński and Demkowicz-Dobrzański, 2013). Derivations of the precision bounds using the above two techniques in the context of optical interferometry are discussed in Sec. VI

For completeness, we list below some other important properties of the QFI. QFI does not increase under a parameter independent quantum channel

$$F_Q(\rho_\varphi) \geq F_Q[\Lambda(\rho_\varphi)], \quad (82)$$

where  $\Lambda$  is an arbitrary completely positive (CP) map. QFI appears in the lowest order expansion of the measure of fidelity  $\mathcal{F}$  between two quantum states

$$\mathcal{F}(\rho_1, \rho_2) = \left( \text{Tr} \sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right)^2, \quad (83)$$

$$\mathcal{F}(\rho_\varphi, \rho_{\varphi+d\varphi}) = 1 - \frac{1}{4} F_Q(\rho_\varphi) d\varphi^2 + O(d\varphi^4). \quad (84)$$

QFI is convex

$$F_Q \left( \sum_i p_i \rho_\varphi^{(i)} \right) \leq \sum_i p_i F_Q(\rho_\varphi^{(i)}), \quad \sum_i p_i = 1, \quad p_i \geq 0, \quad (85)$$

which reflects the fact that mixing quantum states can only reduce achievable estimation sensitivity. In a commonly encountered case, specifically in the context of optical interferometry, when the estimated parameter is encoded on the state by a unitary

$$\rho_\varphi = U_\varphi \rho U_\varphi^\dagger, \quad U_\varphi = e^{-i\hat{H}\varphi}, \quad (86)$$

where  $\hat{H}$  is the generating “Hamiltonian”, the general formula for QFI reads:

$$F_Q(\rho_\varphi) = \sum_{i,j} \frac{2|\langle e_i|\hat{H}|e_j\rangle|^2(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} \quad (87)$$

where  $|e_i\rangle$ ,  $\lambda_i$  form the eigendecomposition of  $\rho$ . Note that in this case the QFI does not depend on the actual value of  $\varphi$ . For the pure state estimation case,  $\rho = |\psi\rangle\langle\psi|$ , the QFI is proportional to the variance of  $\hat{H}$ :

$$F_Q(|\psi_\varphi\rangle) = 4\Delta^2 H = 4(\langle\psi|\hat{H}^2|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle^2) \quad (88)$$

and the QCRB (77) takes a particular appealing form resembling the form of the energy-time uncertainty relation:

$$\Delta^2 \varphi \Delta^2 H \geq \frac{1}{4N}. \quad (89)$$

We conclude the discussion of QFI properties by mentioning a very recent and elegant result proving that for unitary parameter encodings, QFI is proportional to the convex roof of the variance of  $\hat{H}$  (Tóth and Petz, 2013; Yu, 2013):

$$F_Q(\rho_\varphi) = 4 \min_{\{p_i, |\psi^{(i)}\rangle\}} \sum_i p_i \Delta^2 H^{(i)}, \quad (90)$$

where the minimum is performed over all decompositions of  $\rho = \sum_i p_i |\psi^{(i)}\rangle\langle\psi^{(i)}|$ , and  $\Delta^2 H^{(i)}$  denotes the variance of  $\hat{H}$  calculated on  $|\psi^{(i)}\rangle$ .

## 2. Bayesian approach

As the quantum mechanical estimation scenario, with a particular measurement scheme chosen, represents nothing but a probabilistic model with outcome probabilities  $p_\varphi(\mathbf{x})$ , we may also apply the Bayesian techniques described in Sec. IV.A.2. The quantum element of the problem, however, i.e. minimization of the average cost function over the choice of measurements is in general highly non-trivial. Fortunately, provided the problem possesses a particular kind of symmetry it may be solved using the concept of covariant measurements (Holevo, 1982).

In the context of optical interferometry, it is sufficient to consider the unitary parameter encoding case as defined in (86), where the estimated parameter  $\varphi$  will correspond to the phase difference in an interferometer, see Fig. 5 in Sec. V. Let us denote a general POVM measurement as  $M_{\mathbf{x}}$  and the corresponding estimator as  $\tilde{\varphi}$ . Since we need to minimize the average cost over both the measurements and the estimators it is convenient to combine these two elements into one by labeling the POVM elements with the estimated values themselves:  $M_{\tilde{\varphi}} = \int d\mathbf{x} M_{\mathbf{x}} \delta[\tilde{\varphi} - \tilde{\varphi}(\mathbf{x})]$ . The expression for the average cost function, (64), takes the form:

$$\langle C \rangle = \iint \frac{d\varphi}{2\pi} \frac{d\tilde{\varphi}}{2\pi} p(\varphi) \text{Tr}(U_\varphi \rho U_\varphi^\dagger M_{\tilde{\varphi}}) C(\tilde{\varphi}, \varphi). \quad (91)$$

and leaves us with the problem of minimization  $\langle C \rangle$  over a general POVM  $M_{\tilde{\varphi}}$  with standard constraints  $M_{\tilde{\varphi}} \geq 0$ ,  $\int \frac{d\tilde{\varphi}}{2\pi} M_{\tilde{\varphi}} = \mathbb{1}$ . Note that for clarity we use the normalized measure  $\frac{d\varphi}{2\pi}$ .

Provided the problem enjoys the  $\varphi$  shift symmetry, so that  $p(\varphi + \varphi_0) = p(\varphi)$ ,  $C(\tilde{\varphi} + \varphi_0, \varphi + \varphi_0) = C(\tilde{\varphi}, \varphi)$ , it may be shown that one does not loose optimality by restricting the class of POVM measurements to

$$M_{\tilde{\varphi}} = U_{\tilde{\varphi}} \Xi U_{\tilde{\varphi}}^\dagger, \quad (92)$$

which is a special case of the so-called *covariant* measurements (Bartlett *et al.*, 2007; Chiribella *et al.*, 2005; Holevo, 1982). If we take flat prior distribution  $p(\varphi) = 1$ , and the natural cost function  $C(\tilde{\varphi}, \varphi) = 4 \sin^2[(\tilde{\varphi} - \varphi)/2]$  introduced in Sec. IV.A.2, symmetry conditions are fulfilled and substituting (92) into (91) we get a simple expression:

$$\langle C \rangle = \text{Tr} \{ \langle \rho_\varphi \rangle_C \Xi \}, \quad (93)$$

where  $\langle \rho_\varphi \rangle_C = 4 \int \frac{d\varphi}{2\pi} U_\varphi \rho U_\varphi^\dagger \sin^2(\frac{\varphi}{2})$  is the final quantum state averaged with the cost function. Looking for the optimal Bayesian strategy now amounts to minimization of the above quantity over  $\Xi$  with the POVM constraints requiring that  $\Xi \geq 0$ ,  $\int \frac{d\tilde{\varphi}}{2\pi} U_{\tilde{\varphi}} \Xi U_{\tilde{\varphi}}^\dagger = \mathbb{1}$ . As shown in Sec. V and Sec. VI this minimization is indeed possible for optical interferometric estimation models and allows to find the optimal Bayesian strategy and the corresponding minimal average cost.

## V. QUANTUM LIMITS IN DECOHERENCE-FREE INTERFEROMETRY

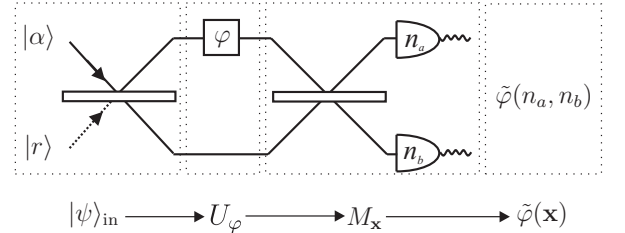


FIG. 5 Instead of a particular Mach-Zehnder interferometric strategy a general quantum interferometric scheme involves a general input probe state  $|\psi\rangle_{\text{in}}$  that is subject to a unitary phase delay operation  $U_\varphi$  followed by a general quantum measurement (POVM)  $M_{\mathbf{x}}$ . Finally, an estimator  $\tilde{\varphi}(\mathbf{x})$  is used to obtain the estimated value of the phase delay.

In order to analyze the ultimate precision bounds in interferometry one needs to employ the quantum estimation theory introduced in Sec. IV.B. In this section we review the most important results of decoherence-free interferometry leaving the analysis of the impact of decoherence for Sec. VI. This will provide us with precision benchmarks to which we will be able to compare realistic

estimation schemes and analyze the reasons for the departure of practically achievable precisions from idealized scenarios.

Formally, interferometry may be regarded as a channel estimation problem where a known state  $|\psi\rangle_{\text{in}}$  is sent through a quantum channel  $U_\varphi = e^{-i\varphi\hat{J}_z}$ , with an unknown parameter  $\varphi$ , where  $\hat{J}_z = \frac{1}{2}(a^\dagger a - b^\dagger b)$  is the  $z$  component of angular momentum operator introduced in Sec. III.A. Quantum measurement  $\hat{M}_{\mathbf{x}}$  is performed at the output state and the value of  $\varphi$  is estimated based on the obtained outcome  $\mathbf{x}$  through an estimator  $\tilde{\varphi}(\mathbf{x})$ , see Fig. 5. Pursuing either QFI or Bayesian approach it is possible to derive bounds on achievable precision that are valid irrespective of how sophisticated is the measurement-estimation strategy employed and how exotic the input states of light are. We start by considering definite photon-number states using both QFI and Bayesian approaches and then move on to discuss issues that arise when discussing fundamental bounds taking into account states of light with indefinite photon number.

#### A. Quantum Fisher Information approach

As discussed in detail in Sec. IV, the QFI approach is particularly well suited to analyze problems where one wants to estimate small deviations of  $\varphi$  around a known value  $\varphi_0$ , as in this local regime the QCRB, (77), is known to be saturable. This is, for example, the case of gravitational-wave interferometry in which one sets the interferometer at the dark fringe and wants to estimate small changes in the interference pattern induced by the passing wave (LIGO Collaboration, 2011, 2013; Pitkin *et al.*, 2011).

Since the state of light at the output,  $|\psi_\varphi\rangle = e^{-i\varphi\hat{J}_z}|\psi\rangle_{\text{in}}$ , is pure, the QFI may be calculated using the simple formula given in Eq. (79). Realizing that  $|\dot{\psi}_\varphi\rangle = -i\hat{J}_z|\psi_\varphi\rangle$  we get the QFI and the corresponding QCRB on the estimation precision:

$$F_Q = 4 \left( \langle \psi_\varphi | \hat{J}_z^2 | \psi_\varphi \rangle - |\langle \psi_\varphi | \hat{J}_z | \psi_\varphi \rangle|^2 \right) = 4\Delta^2 J_z, \Delta\varphi \geq \frac{1}{2\Delta J_z}. \quad (94)$$

Note that the form of the QCRB above may be regarded as an analogue of the Heisenberg uncertainty relation for phase and angular momentum.

Clearly, according to the above bound, the optimal probe states for interferometry are the ones that maximize  $\Delta J_z$ . We fix the total number of photons and look for  $N$ -photon states maximizing  $\Delta J_z$ . A General  $N$ -photon input state and the explicit form of the corre-

sponding QFI read:

$$|\psi\rangle_{\text{in}} = \sum_{n=0}^N c_n |n\rangle |N-n\rangle, \quad (95)$$

$$F_Q = 4 \left[ \sum_{n=0}^N |c_n|^2 n^2 - \left( \sum_{n=0}^N |c_n|^2 n \right)^2 \right]. \quad (96)$$

Let us first consider the situation in which  $|\psi\rangle_{\text{in}}$  is a state resulting from sending  $N$  photons on the balanced beam-splitter, as discussed in Sec. III.C. This time, however, we do not insist on a particular measurement nor estimation scheme, but just want to calculate the corresponding QCRB on the sensitivity. Written in the photon-number basis the state takes the form:

$$|\psi\rangle_{\text{in}} = \sum_{n=0}^N \sqrt{\frac{1}{2^N} \binom{N}{n}} |n\rangle |N-n\rangle, \quad (97)$$

for which  $F_Q = N$  and this results in the shot noise bound on precision:

$$\Delta\varphi^{|N\rangle|0\rangle} \geq \frac{1}{\sqrt{N}}. \quad (98)$$

Recall, that this bound is saturated with the simple MZ interferometric scheme discussed in Eq. (36), which is a proof that for the considered probe state this measurement-estimation scheme is optimal. From a particle-description point of view, see Sec. II.D, the above considered state is a pure product state with no entanglement between the photons. More generally, the shot noise limit, sometimes referred to as the standard quantum limit, is valid for all  $N$ -photon separable states (Giovannetti *et al.*, 2011; Pezzé and Smerzi, 2009), and going beyond this limit requires making use of inter-photon entanglement, see Sec. V.D.

Let us now investigate general input states, which possibly may be entangled. Consider the state of the form  $|\text{NOON}\rangle = \frac{1}{\sqrt{2}}(|N\rangle|0\rangle + |0\rangle|N\rangle)$  which is commonly referred to as the “NOON” state (Bollinger *et al.*, 1996; Lee *et al.*, 2002). QFI for such state is given by  $F_Q = N^2$  and consequently

$$\Delta\varphi \geq \frac{1}{N}, \quad (99)$$

which is referred to as the *Heisenberg limit*. In fact, NOON state gives the best possible precision as it has the biggest variance of  $\hat{J}_z$  among the states with a given photon number  $N$  (Bollinger *et al.*, 1996; Giovannetti *et al.*, 2006). Still, the practical usefulness of the NOON states is doubtful. The difficulty in their preparation increases dramatically with increasing  $N$ , and with present technology the experiments are limited to relatively small  $N$ , e.g.  $N = 4$  (Nagata *et al.*, 2007) or  $N = 5$  (Afek *et al.*,



2010). Moreover, even if prepared, their extreme susceptibility to decoherence with increasing  $N$ , see Sec. VI, makes them hardly useful in any realistic scenario unless  $N$  is restricted to small values. Taking into account that experimentally accessible squeezed states of light offer a comparable performance in the decoherence-free scenario, see Sec. V.C, and basically optimal asymptotic performance in the presence of decoherence, see Sec. VI, there is not much in favor of the NOON states apart from their conceptual appeal.

## B. Bayesian approach

Let us now look for the fundamental precision bounds in the Bayesian approach (Berry and Wiseman, 2000; Hradil *et al.*, 1996). We assume the flat prior distribution  $p(\varphi) = 1/2\pi$  reflecting our complete initial ignorance on the true phase value, and the natural cost function  $C(\tilde{\varphi}, \varphi) = 4\sin^2[(\tilde{\varphi} - \varphi)/2]$ , see Sec. IV.A.2. Thanks to the phase shift symmetry of the problem, see Sec. IV.B.2, we can restrict the class of measurements to covariant measurements  $M_{\tilde{\varphi}} = U_{\tilde{\varphi}}\Xi U_{\tilde{\varphi}}^\dagger$ , where  $\Xi$  is the seed measurement operator, to be optimized below. Using Eq. (93), and noting that  $\int_0^{2\pi} \frac{d\varphi}{2\pi} 4\sin^2(\varphi/2)e^{i\varphi(n-m)} = 2\delta_{nm} - (\delta_{n,m-1} + \delta_{n,m+1})$ , the averaged cost for a general  $N$ -photon input state  $|\psi\rangle = \sum_{n=0}^N c_n |n, N-n\rangle$  reads:

$$\langle C \rangle = 2 - 2\text{Re} \left( \sum_{n=1}^N c_n^* c_{n-1} \Xi_{n,n-1} \right). \quad (100)$$

Because  $\Xi$  is Hermitian, the completeness condition  $\int \frac{d\varphi}{2\pi} \hat{U}_\varphi \Xi \hat{U}_\varphi^\dagger = \mathbb{1}$  implies that  $\Xi_{nn} = 1$ , while due to the positive semi-definiteness condition  $\Xi \geq 0$ ,  $|\Xi_{nm}| \leq \sqrt{\Xi_{nn}\Xi_{mm}} = 1$ . Therefore, The real part in the subtracted term in Eq. (100) can at most be  $\sum_{n=1}^N |c_n| |c_{n-1}|$ , which will be the case for  $\Xi_{n,m} = e^{i(\xi_n - \xi_m)}$ , where  $\xi_n = \arg(c_n)$ . This is a legitimate positive semi-definite operator, as it can be written as  $\Xi = |e_N\rangle\langle e_N|$ , with  $|e_N\rangle = \sum_{n=0}^N e^{i\xi_n} |n, N-n\rangle$  (Chiribella *et al.*, 2005; Holevo, 1982). Thus, for a given input state the optimal Bayesian measurement-estimation strategy yields

$$\langle C \rangle = 2 \left( 1 - \sum_{n=1}^N |c_n c_{n-1}| \right). \quad (101)$$

For the uncorrelated input state (97), the average cost reads:

$$\langle C \rangle = 2 \left( 1 - \frac{1}{2^N} \sum_{n=0}^{N-1} \sqrt{\binom{N}{n} \binom{N}{n+1}} \right) \stackrel{N \rightarrow \infty}{\approx} \frac{1}{N}. \quad (102)$$

Since in the limit of small estimation uncertainty the considered cost function approaches the MSE, we may conclude that in the limit of large  $N$ :

$$\Delta\varphi \stackrel{N \rightarrow \infty}{\approx} \frac{1}{\sqrt{N}}, \quad (103)$$

which coincides with the standard shot-noise limit derived within the QFI approach.

Note a subtle difference between the above solution and the solution of the optimal Bayesian transmission coefficient estimation problem discussed in Sec. IV.A.3 with the  $\varphi$  parametrization employed. The formulas for probabilities in Sec. IV.A.3, can be regarded as arising from measuring each of uncorrelated photons leaving the interferometer independently, while in the present considerations we have allowed for arbitrary quantum measurements, which are in general collective. Importantly, we account for the adaptive protocols in which a measurement on a subsequent photon depends on the results obtained previously (Kołodziej and Demkowicz-Dobrzański, 2010)—practically these are usually additional controlled phase shifts allowing to keep the setup at the optimal operation point (Higgins *et al.*, 2007). This approach is therefore more general and in particular, does not suffer from a  $\pm\varphi$  ambiguity that forced us to restrict the estimated region to  $[0, \pi)$  in order to obtain nontrivial results given in Eq. (75).

Let us now look for the optimal input states. From Eq. (102) it is clear that we may restrict ourselves to real  $c_n$ . Denoting by  $\mathbf{c}$  the vector containing the state coefficients  $c_n$ , we rewrite the formula for  $\langle C \rangle$  in a more appealing form

$$\langle C \rangle = 2 - \mathbf{c}^T \mathbf{A} \mathbf{c}, \quad A_{n,n-1} = A_{n-1,n} = 1, \quad (104)$$

from which it is clear that minimizing the cost function is equivalent to finding the eigenvector with maximal eigenvalue of the matrix  $\mathbf{A}$ , which has all its entries zero except for its first off-diagonals. This can be done analytically (Berry and Wiseman, 2000; Luis and Perina, 1996; Summy and Pegg, 1990) and the optimal state, which we will refer to as the *sine* state, together with the resulting cost read

$$|\psi\rangle_{\text{in}} = \sum_{n=0}^N \sqrt{\frac{2}{2+N}} \sin\left(\frac{n+1}{N+2}\pi\right) |n\rangle |N-n\rangle, \quad (105)$$

$$\langle C \rangle = 2 \left[ 1 - \cos\left(\frac{\pi}{N+2}\right) \right] \stackrel{N \rightarrow \infty}{\approx} \frac{\pi^2}{N^2}. \quad (106)$$

Again, in the large  $N$  limit we may identify the average cost with the average MSE, so that the asymptotic precision reads:

$$\Delta\varphi \stackrel{N \rightarrow \infty}{\approx} \frac{\pi}{N}. \quad (107)$$

Analogously, as in the QFI approach we arrive at the  $1/N$  Heisenberg scaling of precision, but with an additional constant factor  $\pi$ , reflecting the fact that Bayesian approach is more demanding as it requires the strategy to work well under complete prior ignorance of the value of the estimated phase. Note also that the structure of the optimal states is completely different from the NOON

states. In fact, the NOON states are useless in absence of any prior knowledge on the phase, since they are invariant under  $2\pi/N$  phase shift, and hence cannot unambiguously resolve phases differing by this amount.

### C. Indefinite photon-number states

We now consider a more general class of states with indefinite photon numbers and look for optimal probe states treating the *average* photon number  $\langle N \rangle$  as a fixed resource. A state with an indefinite number of photons may possess in general coherences between sectors with different total numbers of particles. These coherences may in principle improve estimation precision. However, a photon number measurement performed at the output ports projects the state on one of the sectors and necessarily all coherences between different total photon number sectors are destroyed. In order to benefit from these coherences, one needs to make use of a more general scheme such as e.g. homodyne detection, where an additional phase reference beam is needed, typically called the local oscillator, which is being mixed with the signal light at the detection stage. Usually, the local oscillator is assumed to be strong, classical field with a well defined phase. In other words, it provides one with reference frame with respect to which phase of the signal beams can be measured (Bartlett *et al.*, 2007; Mølmer, 1997). Thus, it is crucial to explicitly state whether the reference beam is included in the overall energy budget or is it treated as a free resource. Otherwise one may arrive at conflicting statements on the achievable fundamental bounds (Jarzyna and Demkowicz-Dobrzański, 2012).

#### 1. Role of the reference beam

As an illustrative example, consider an artificial one mode scheme with input in coherent state  $|\psi\rangle = |\alpha\rangle$  which passes through the phase delay  $\varphi$ . Strictly speaking, this is not an interferometer and one may wonder how one can possibly get any information on the phase by measuring the output state. Still  $|\psi\rangle$  evolves into a formally different state  $|\psi_\varphi\rangle = U_\varphi|\alpha\rangle = |\alpha e^{-i\varphi}\rangle$ , where  $U_\varphi = e^{-i\varphi\hat{n}}$ , and since the corresponding QFI is non-zero

$$F_Q = 4(\langle\alpha|\hat{n}^2|\alpha\rangle - \langle\alpha|\hat{n}|\alpha\rangle^2) = 4|\alpha|^2, \quad (108)$$

it is in principle possible to draw some information on the phase by measuring  $|\psi_\varphi\rangle$ . Clearly, the measurement required cannot be a direct photon-number measurement, and an additional phase reference beam needs to be mixed with the state before sending the light to the detectors. In a fair approach one should include the reference beam into the setup and assume that whole state of signal+reference beam is averaged over a global undefined phase. This formalizes the notion of the relative

phase delay - it is defined with respect to reference beam and there is no such thing like absolute phase delay.

More formally, let  $|\psi\rangle_{ar} = |\alpha\rangle_a|\beta\rangle_r$ , be the original coherent state used for sensing the phase, accompanied by a coherent reference beam  $|\beta\rangle$ . The corresponding output state reads  $|\psi_\varphi\rangle_{ar} = |\alpha e^{-i\varphi}\rangle_a|\beta\rangle_r$ . Now, the phase  $\varphi$  plays the role of the *relative* phase shift between the two modes, with a clear physical interpretation. The combined phase shift in the two modes, i.e. an operation  $U_\theta^a U_\theta^r = e^{-i\theta(\hat{n}_a + \hat{n}_r)}$  has no physical significance as it is not detectable without an ... additional reference beam. Hence, before calculating the QFI or any other quantity determining fundamental precision bounds, one should first average the state  $|\psi\rangle_{ar}$  over the combined phase shift and treat the resulting density matrix as the input probe state

$$\rho = \int_0^{2\pi} \frac{d\theta}{2\pi} U_\theta^a U_\theta^r |\psi\rangle_{ar} \langle\psi| \otimes |\beta\rangle\langle\beta| U_\theta^{a\dagger} U_\theta^{r\dagger} = \sum_{N=0}^{\infty} p_N \rho_N. \quad (109)$$

This operation destroys all the coherences between sectors with different total photon number, and the resulting state is a mixture of states  $\rho_N$  with different total photon numbers  $N$ , appearing with probabilities  $p_N$ .

In the absence of a reference beam, i.e. when  $\beta = 0$ , the above averaging kills all the coherences between terms with different photon numbers in the mode  $a$ :

$$\begin{aligned} \rho &= \int \frac{d\theta}{2\pi} |\alpha e^{-i\theta}\rangle \langle\alpha e^{-i\theta}| \otimes |0\rangle\langle 0| = \\ &= e^{-|\alpha|^2} \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} |n\rangle\langle n| \otimes |0\rangle\langle 0| \end{aligned} \quad (110)$$

which results in a state insensitive to phase delays and gives  $F = 0$ , restoring our physical intuition. On the other hand,  $F = 4|\alpha|^2$  obtained previously is recovered in the limit  $|\beta|^2 \rightarrow \infty$ , meaning that reference beam is classical—consists of many more photons than the signal beams.

The above averaging prescription, is valid in general also when label  $a$  refers to more than one mode. Considering the standard Mach-Zehnder interferometer fed with a state  $|\psi\rangle_{ab}$  with an indefinite photon numbers and no additional reference beam available one again needs to perform the averaging over a common phase shift. If this is not done, one may obtain conflicting results on e.g. QFI for seemingly equivalent phase shift operations such as  $U_\varphi = e^{-i\hat{n}_a\varphi}$  or  $U'_\varphi = e^{-i(\hat{n}_a - \hat{n}_b)\varphi/2}$ . The reason is that, without the common phase averaging, one implicitly assumes the existence of a strong external classical phase reference to which the phase shifts are defined. In particular,  $U_\varphi$  assumes that the second mode is perfectly locked with the external reference beam and is not affected by the phase shift, whereas  $U'_\varphi$  assumes that there are exactly opposite phase shifts in the two modes with respect to the reference. Different choices

of „phase-shift distribution” may lead to a factor of 2 or even factor of 4 discrepancies in the reported QFIs in apparently equivalent optical phase estimation schemes—see (Jarzyna and Demkowicz-Dobrzański, 2012) for further discussion and compare with some of the results that were obtained without the averaging performed (Joo *et al.*, 2011; Spagnolo *et al.*, 2012).

One can also understand why ignoring the need for a reference beam may result in underestimating the required energy resources. Consider a single mode state with an indefinite photon number  $|\psi\rangle = \sum_{n=0}^N c_n |n\rangle$  evolving under  $e^{-i\varphi\hat{n}}$  and note that, from the phase sensing point of view, this situation is formally equivalent to a two-mode state with a definite photon number  $N$ :  $|\psi\rangle_{ab} = \sum_{n=0}^N c_n |n\rangle |N-n\rangle$ , evolving under  $e^{-i\varphi\hat{n}_a}$ . Still, the average photon number consumed in the one mode case is  $\langle N \rangle = \sum_{n=0}^N |c_n|^2 n$  and is in general smaller than  $N$ .

## 2. Optimal indefinite photon number strategies

Looking for the optimal states with fixed *average* photon number is in general more difficult than in the definite photon-number case. Still, if we agree with the above-advocated approach to average all the input states over a common phase-shift transformation as in Eq. (109), then the resulting state is a probabilistic mixture of definite photon number states. Intuitively, it is then clear that, instead of sending the considered averaged state, it is more advantageous to have information which particular component  $\rho_N$  of the mixture is being sent. This would allow to adjust the measurement-estimation procedure to a given component and improve the overall performance. This intuition is reflected by the properties of both the QFI and the Bayesian cost, which are respectively convex and concave quantities (Helstrom, 1976):

$$F_Q \left( \sum_N p_N \rho_N \right) \leq \sum_N p_N F_Q(\rho_N), \quad (111)$$

$$\left\langle C \left( \sum_N p_N \rho_N \right) \right\rangle \geq \sum_N p_N \langle C(\rho_N) \rangle. \quad (112)$$

This, however, implies that knowing the solution for the optimal *definite* photon number probe states, by adjusting the probabilities  $p_N$  with which different optimal  $\rho_N$  are being sent, we may determine the optimal strategies with the average photon-number fixed.

Taking the QFI approach for a moment, we recall that the optimal  $N$ -photon state, the NOON state, yields  $F_Q(\rho_N) = N^2$ . Let us consider a strategy where a vacuum state and the NOON state are sent with probabilities  $1-p$  and  $p$  respectively, with the constraint on the average photon number  $pN = \langle N \rangle$ . The corresponding QFI reads  $F_Q = (1-p) \cdot 0 + pN^2$ . Substituting  $p = \langle N \rangle / N$

we get  $F_Q = N \langle N \rangle$ . Therefore, while keeping  $\langle N \rangle$  fixed we may increase  $N$  arbitrarily and in principle reach  $F_Q = \infty$ , suggesting the possibility of arbitrary good sensing precisions (Rivas and Luis, 2012; Zhang *et al.*, 2013). Note in particular, that a naive generalization of the Heisenberg limit to  $\Delta\varphi \geq \frac{1}{\langle N \rangle}$ , does not hold, and the strategies beating this bound are referred to as sub-Heisenberg strategies (Anisimov *et al.*, 2010). A universally valid bound may be written as  $\Delta\varphi \geq 1/\sqrt{\langle \hat{N}^2 \rangle}$  (Hofmann, 2009), but the question remains, whether the bound is saturable, and in particular does quantum mechanics indeed allows for practically useful estimation protocols leading to the sub-Heisenberg scaling of precision. Closer investigations of that problem proves such hypothesis to be false (Berry *et al.*, 2012; Giovannetti and Maccone, 2012; Tsang, 2012; Zwiernik *et al.*, 2010). In principle we may achieve the sub-Heisenberg precision in the local estimation regime but in order for the local strategy to be valid, we should know the value of estimated parameter with prior precision of the same order as the one we want to obtain, what makes the utility of the procedure questionable. Actually, if no such assumption on the priori knowledge is made, the Heisenberg scaling in the form  $\Delta\varphi = \text{const}/\langle N \rangle$  is the best possible scaling of precision.

This claim can also be confirmed within the Bayesian approach with flat prior phase distribution. For large  $N$  the minimal Bayesian cost behaves like  $\langle C(\rho_N) \rangle = \pi^2/N^2$ , see Eq. (105). Since this function is convex, taking convex combinations of the cost for two different total photon numbers  $N_1, N_2$ , such that  $p_1 N_1 + p_2 N_2 = \langle N \rangle$  will yield the cost higher than  $\pi^2/\langle N \rangle^2$ , and the corresponding uncertainty  $\Delta\varphi \geq \pi/\langle N \rangle$ , indicating the universal validity of the Heisenberg scaling of precision.

## 3. Gaussian states

From a practical point of view, rather than looking for the optimal indefinite photon number states for interferometry it is more important to analyze experimentally accessible Gaussian states. The paradigmatic example of a Gaussian state applied in quantum enhanced interferometry is the two mode state  $|\alpha\rangle|r\rangle$ —coherent state in mode  $a$  and squeezed vacuum in mode  $b$ . We have already discussed this example in Sec. III.D, and calculated the precision for a simple measurement-estimation scheme. For such states, sent through a fifty-fifty beam splitter, quantum Fisher information can be calculated explicitly (Jarzyna and Demkowicz-Dobrzański, 2012; Ono and Hofmann, 2010; Pezzé and Smerzi, 2008):

$$F_Q = |\alpha|^2 e^{2r} + \sinh^2 r. \quad (113)$$

For the extreme cases  $|\alpha|^2 = 0, \sinh^2 r = \langle N \rangle$  and  $|\alpha|^2 = \langle N \rangle, \sinh^2 r = 0$  this formula gives  $F_Q = \langle N \rangle$ ,

implying the shot noise scaling. Most importantly, optimization of Eq. (113) over  $\alpha$  and  $r$  with constraint  $|\alpha|^2 + \sinh^2 r = \langle N \rangle$  gives asymptotically the Heisenberg limit  $\Delta\varphi \sim 1/\langle N \rangle$ , making this strategy as good as the NOON-one for large number of photons. Moreover, this bound on precision can be saturated by estimation strategies based on photon-number (Pezzé and Smerzi, 2008; Seshadreesan *et al.*, 2011) or homodyne (D'Ariano *et al.*, 1995) measurements. This also proves that a simple measurement-estimation strategy discussed in Sec. III.D which yielded  $1/\langle N \rangle^{3/4}$  scaling of precision is not optimal. Unlike the simple interferometric scheme where it was optimal to dedicate approximately  $\sqrt{\langle N \rangle}$  photons to the squeezed beam, from the QFI point of view it is optimal to equally divide the number of photons between the coherent and the squeezed beam.

More generally, finding the fundamental limit on precision achievable with general Gaussian states, requires optimization of the QFI or the Bayesian average cost function over general two-mode Gaussian input states, specified by the covariance matrix and the vector of first moments, see Sec. II.B. For the decoherence-free case this was done in Pinel *et al.* (2012, 2013). Crucial observation is that for pure states, the overlap between two M-mode Gaussian states  $|\psi_\varphi\rangle$  and  $|\psi_\varphi + d\varphi\rangle$  is given by (up to the second order in  $d\varphi$ )

$$|\langle\psi_\varphi|\psi_{\varphi+d\varphi}\rangle|^2 = 1 - \frac{d\varphi^2}{4} \left( 2(4\pi)^2 \int \left( \frac{dW(\mathbf{z})}{d\varphi} \right)^2 d^{2M}\mathbf{z} \right) \quad (114)$$

where  $W(\mathbf{x})$  is the Wigner function (5) of state  $|\psi_\varphi\rangle$ . Thus, because  $|\langle\psi_\varphi|\psi_{\varphi+d\varphi}\rangle|^2 = 1 - \frac{1}{4}F_Q d\varphi^2$  we may write that

$$\Delta\varphi \geq \left( 2(4\pi)^2 \int \left( \frac{dW(\mathbf{z})}{d\varphi} \right)^2 d^{2M}\mathbf{z} \right)^{-\frac{1}{2}}. \quad (115)$$

In terms of the covariance matrix  $\sigma$  and the first moments  $\langle \mathbf{z} \rangle$  the formula takes an explicit form:

$$\Delta\varphi \geq \left( \frac{d\langle \mathbf{z} \rangle}{d\varphi}^T \sigma^{-1} \frac{d\langle \mathbf{z} \rangle}{d\varphi} + \frac{1}{4} \text{Tr} \left( \left( \frac{d\sigma}{d\varphi} \sigma^{-1} \right)^2 \right) \right)^{-\frac{1}{2}}. \quad (116)$$

Formal optimization of the above equation was done by Pinel *et al.* (2012), however, the result was a one-mode squeezed-vacuum state, which in order to carry phase information needs to be assisted by a reference beam. Unfortunately, performing a common phase-averaging procedure described in Sec. V.C.1 in order to calculate the precision in the absence of additional phase reference destroys the Gaussian structure of the state and makes the optimization intractable. Luckily, for path symmetric states, i.e. the states invariant under the exchange of interferometer arms, the phase averaging procedure does not affect the QFI (Jarzyna and Demkowicz-Dobrzański,

2012). Hence, assuming the path-symmetry the optimal Gaussian state is given by  $|r\rangle|r\rangle$  with  $\sinh^2 r = \langle N \rangle/2$ —two squeezed vacuums send into the input ports of the interferometer—and its corresponding QFI leads to the QCRB

$$\Delta\varphi \geq \frac{1}{\sqrt{\langle N \rangle(\langle N \rangle + 2)}} \approx \frac{1}{\langle N \rangle} \quad (117)$$

The state also achieves the Heisenberg limit for a large number of photons in the setup but it does not require any external phase reference. It is also worth noting that this state, while being mode-separable is particle-entangled and is feasible to prepare with current technology for moderate squeezing strengths. However, the enhancement over optimal squeezed-coherent strategy is rather small and vanish for large number of photons. Precisions obtained in squeezed-coherent and squeezed-squeezed scenarios are depicted in Fig. 6.

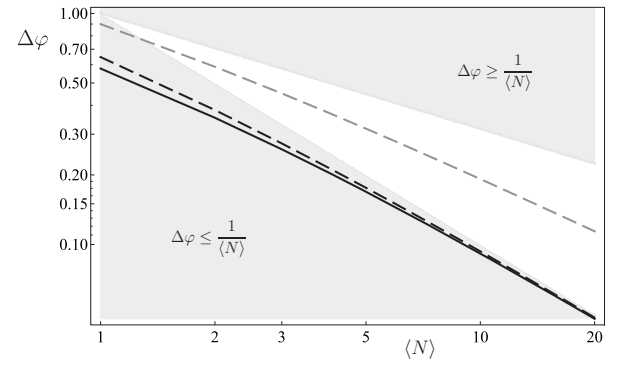


FIG. 6 Limits on precision obtained within QFI approach when using two optimally squeezed states in both modes  $|r\rangle|r\rangle$  (black, solid), coherent and squeezed-vacuum states  $|\alpha\rangle|r\rangle$  (black, dashed). For comparison precision achievable with simple coherent and squeezed vacuum MZ interferometric scheme discussed in Sec. III.D is also depicted (gray, dashed).

One can also study Gaussian states within the Bayesian framework. Optimal seed operator can be easily generalized from the definite photon number case to  $\Xi = \sum_{N=0}^{\infty} |e_N\rangle\langle e_N|$ . Conceptually, the whole treatment is the same as in the definite photon number case. However, the expressions and calculations are very involved and will not be presented here.

#### D. Role of entanglement

The issue of entanglement is crucial in quantum interferometry as it is known that only entangled states can beat the shot noise scaling (Pezzé and Smerzi, 2009). This statement is sometimes questioned, pointing out the example of the squeezed+coherent light strategy, where the interferometer is fed with seemingly unentangled  $|\alpha\rangle|r\rangle$  input state. The reason of confusion is the



conflict of notions of mode and particle entanglement. As discussed in detail in Sec. II.E, the two notions are not compatible, and there are states which are particle entangled, while having no mode entanglement and vice versa. In the context of interferometry it is the *particle* entanglement that is the source of quantum enhancement of precision. In order to avoid criticism based on the ground of fundamental indistinguishability of particles and therefore a questionable physical content of the distinguishable particle-based entanglement picture on the fundamental level (Benatti *et al.*, 2010), we should stress that when considering models involving indistinguishable particles one should regard this statement as a formal (but still a meaningful and useful) criterion where the particles are treated as formally distinguishable as described in Sec. II.D.

To see this, let us consider first a separable input state of  $N$  photons of the form  $\rho = \rho_1 \otimes \dots \otimes \rho_N$ , where  $\rho_i$  denotes the state of the  $i$ -th photon. Since the phase shift evolution affects each of the photons independently  $\rho_\varphi = U_\varphi^{\otimes N} \rho U_\varphi^{\otimes N\dagger}$  and the QFI is additive on product states we may write:

$$F_Q(\rho) = \sum_{i=1}^N F_Q(\rho_i) \leq N F_Q(\rho_{\max}) \quad (118)$$

where  $\rho_{\max}$  denotes state from the set  $\{\rho_i\}_{i=1,\dots,N}$  for which QFI takes the largest value. But for a one photon state, the maximum value of QFI is equal to 1, so

$$F(\rho) \leq N, \quad \Delta\varphi \geq \frac{1}{\sqrt{N}}. \quad (119)$$

For general separable states  $\rho = \sum_i p_i \rho_1^{(i)} \otimes \dots \otimes \rho_N^{(i)}$  it is sufficient to use the convexity of QFI, together with Eq. (119) to obtain the same conclusion. Above results imply that QFI, or precision, can be interpreted as a particle-entanglement witness, i.e. all states that give precision scaling better than the shot noise must be particle-entangled (Hyllus *et al.*, 2012; Tóth, 2012).

The seemingly unentangled state  $|\alpha\rangle|r\rangle$  when projected on the definite total photon number sector, indeed contains particle entanglement as was demonstrated in Sec. II.E. This fact should be viewed as the fundamental source of its ability for performing quantum-enhanced sensing. It is also worth stressing, that unlike mode entanglement, particle entanglement is invariant under passive optical transformation like beam splitters, delay lines and mirrors, which makes it a sensible quantity to be treated as a resource for quantum enhanced interferometry.

### E. Multi-pass protocols

A common method, used in e.g. gravitational wave detectors, to improve interferometric precision is to let

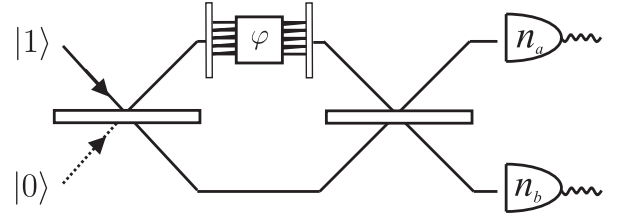


FIG. 7 A multi-pass interferometric protocol. A standard phase delay element is replaced by a setup which makes the beam to pass through the phase delay multiple number of times.

the light bounce back and forth through the phase delay element many times so that the phase delay signal is enhanced as shown in Fig. 7. This method is used in GEO600 experiment (LIGO Collaboration, 2011), where the light bounces twice in each of the sensing arms, making the detector as sensitive as the one with arms twice as long. Up to some approximation, one can also view the Fabry-Perot cavities placed on top of the Mach-Zehnder design as devices forcing each of the photon to pass multiple-times through the arms of the interferometer and acquire a multiple of the phase delay (Berry *et al.*, 2009; Demkowicz-Dobrzański *et al.*, 2013).

Consider a single photon in the state (after the first beam splitter)  $|\psi\rangle_{\text{in}} = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ . After passing through the phase shift  $N$  times it evolves into  $|\psi_\varphi\rangle = \frac{1}{\sqrt{2}}(|01\rangle + e^{-iN\varphi}|10\rangle)$ . The phase is acquired  $N$  times faster compared with a single pass case, mimicking the behavior of a single pass experiment with a NOON state. Hence, the precision may in principle be improved by a factor of  $N$ . Treating the number of single photon passes as a resource, it has been demonstrated experimentally (Higgins *et al.*, 2007) that in the absence of noise such a device can indeed achieve the Heisenberg scaling without resorting to entanglement and efficiency of various multi-pass protocols has been analyzed in detail in (Berry *et al.*, 2009). This is not to say, that all quantum strategies are formally equivalent to single-photon multi-pass strategies. As will be discussed in the next section, the NOON states are extremely susceptible to decoherence, in particular loss, and this property is shared by the multi-pass strategies. Other quantum strategies prove more advantageous in this case, and they do not have a simple multi-pass equivalent (Demkowicz-Dobrzański, 2010; Kaftal and Demkowicz-Dobrzański, 2014).

## VI. QUANTUM LIMITS IN REALISTIC INTERFEROMETRY

In this section we revisit the ultimate limits on precision derived in Sec. V taking into account realistic noise effects. We study three decoherence processes that are typically taken into account when discussing imperfec-

tions in interferometric setups. We consider the effects of *phase diffusion*, *photonic losses* and the impact of *imperfect visibility*, see Fig. 8. In order to establish the ultimate limits on the estimation performance, we first analyze the above three decoherence models using the QFI perspective and secondly compare the bounds obtained with the ones derived within the Bayesian approach.

For the most part of this section, we will consider input states with definite number of photons,  $N$ , so that

$$\rho_{\text{in}} = |\psi\rangle_{\text{in}}\langle\psi|, \quad |\psi\rangle_{\text{in}} = \sum_{n=0}^N c_n |n\rangle |N-n\rangle. \quad (120)$$

Similarly as in Sec. V, this will again be sufficient to draw conclusions also on the performance of indefinite photon number states, which will be discussed in detail in Sec. VI.D.1.

In what follows it will sometimes prove useful to switch from mode to particle description, see Sec. II.D, and treat photons formally as distinguishable particles but prepared in a symmetrized state. This approach is illustrated in Fig. 9 where each photon is represented by a different horizontal line, and travels through a phase encoding transformation  $U_\varphi^{(i)} = e^{-i\varphi\hat{\sigma}_z^{(i)}/2}$ , where  $\hat{\sigma}_z^{(i)}$  is a  $z$  Pauli matrix acting on the  $i$ -th qubit. The combined phase encoding operation is a simple tensor product  $U_\varphi^{\otimes N} = e^{-i\varphi\sum_i \hat{\sigma}_z^{(i)}/2} = e^{-i\varphi\hat{J}_z}$ , recovering the familiar formula but with  $\hat{J}_z$  being now interpreted as the  $z$  component of the total angular momentum which is the sum of individual angular momenta. The photons are then subject to decoherence that acts in either correlated or uncorrelated manner. In the case of *phase diffusion* (i) the decoherence has a collective character since each of the photons experiences the same fluctuation of the phase being sensed, while in the case of *loss* (ii) and *imperfect visibility* (iii) the decoherence map has a tensor structure  $\Lambda^{\otimes N}$  reflecting the fact that it affects each photon independently. In the latter case of independent decoherence models the overall state evolution is uncorrelated and may be written as:

$$\rho_\varphi = \Lambda_\varphi^{\otimes N}(\rho_{\text{in}}), \quad \Lambda_\varphi(\cdot) = \Lambda(U_\varphi \cdot U_\varphi^\dagger). \quad (121)$$

### A. Decoherence models

In general, decoherence is a consequence of the uncontrolled interactions of a quantum system with the environment. Provided the system is initially decoupled from the environment, the general evolution of a quantum system interacting with the environment mathematically corresponds to a *completely positive trace preserving* map  $\Lambda$ . Every  $\Lambda$  can be written using the Kraus representation (Nielsen and Chuang, 2000):

$$\rho_{\text{out}} = \Lambda(\rho_{\text{in}}) = \sum_i K_i \rho_{\text{in}} K_i^\dagger, \quad \sum_i K_i^\dagger K_i = \mathbb{1}, \quad (122)$$

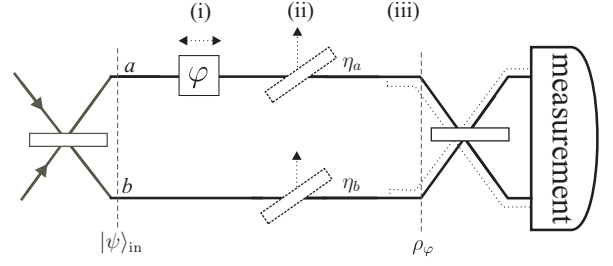


FIG. 8 Schematic description of the decoherence processes discussed that affect the performance of an optical interferometer: (i) *phase diffusion* representing stochastic fluctuations of the estimated phase delay, (ii) *losses* in the respective  $a/b$  arms represented by fictitious beam splitters with  $0 \leq \eta_{a/b} \leq 1$  transmission coefficients, (iii) *imperfect visibility* indicated by a mode mismatch of the beams interfering at the output beam splitter

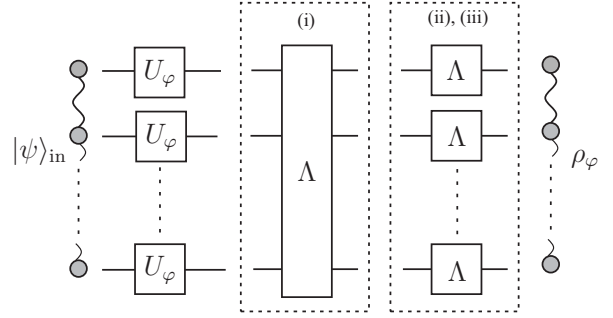


FIG. 9 General metrological scheme in case of photons being treated as formally distinguishable particles. Each photon travels through a phase encoding transformation  $U_\varphi$ . Apart from that all photons are subject to either correlated (i) (phase diffusion), or uncorrelated (ii), (iii) (loss, imperfect visibility) decoherence process.

where  $K_i$  are called the Kraus operators.

Effects of decoherence inside an interferometer are taken into account by replacing the unitary transformation  $U_\varphi$  describing the action of the ideal interferometer, see Sec. V, with its noisy variant  $\Lambda_\varphi$ :

$$\rho_\varphi = \Lambda_\varphi(\rho_{\text{in}}) = \left( \sum_i K_i U_\varphi \rho_{\text{in}} U_\varphi^\dagger K_i^\dagger \right). \quad (123)$$

The formal structure of the above formula corresponds to a situation in which decoherence happens *after* the unitary phase encoding. This of course might not be true in general. Still, for all the models considered in this review the decoherence part and the unitary part commute and therefore the order in which they are written is a matter of convenience.

## 1. Phase diffusion

*Phase diffusion*, also termed as the *collective dephasing* or the *phase noise* represents the effect of fluctuation of the estimated phase delay  $\varphi$ . Such effect may be caused by any process that stochastically varies the effective optical lengths traveled by the photons, such as thermal deformations or the micro-motions of the optical elements. We model the optical interferometry in the presence of phase diffusion process by the following map:

$$\rho_\varphi = \Lambda_\varphi(\rho_{\text{in}}) = \int_{-\infty}^{\infty} d\phi p_\varphi(\phi) U_\phi \rho_{\text{in}} U_\phi^\dagger \quad (124)$$

where the phase delay is a random variable  $\phi$  distributed according to probability distribution  $p_\varphi(\phi)$ . Note that the above form is actually the Kraus representation of the map  $\Lambda_\varphi$  with Kraus operators  $K_\phi = \sqrt{p_\varphi(\phi)} U_\phi$ . In case  $p_\varphi(\phi)$  is a Gaussian distribution with variance  $\Gamma$  and the mean equal to the estimated parameter  $\varphi$ ,  $p_\varphi(\phi) = \frac{1}{\sqrt{2\pi\Gamma}} e^{-\frac{(\phi-\varphi)^2}{2\Gamma}}$ , the output state reads explicitly (Genoni *et al.*, 2011):

$$\rho_\varphi = \sum_{n,m=0}^N c_n c_m^* e^{-\frac{\Gamma}{2}(n-m)^2} e^{-i(n-m)\varphi} |n, N-n\rangle \langle m, N-m|, \quad (125)$$

where  $c_n$  are parameters of the input state given as in (120). The above equation indicates that due to the phase diffusion the off-diagonal elements of  $\rho_\varphi$  are exponentially suppressed at a rate increasing in the anti-diagonal directions.

## 2. Photonic losses

In the *lossy interferometer* scenario, the fictitious beam-splitters introduced in the interferometer arms with respective power transmission coefficient  $\eta_{a/b}$  account for the probability of photons to leak out. Such a loss model is relatively general, as due to the commutativity of the noise with the phase accumulation (Demkowicz-Dobrzanski *et al.*, 2009), it accounts for the photonic losses happening at any stage of the phase sensing process. Moreover, losses at the detection as well as the preparation stages can be moved inside the interferometer provided they are equal in both arms. This makes the model applicable in typical experimental realization of quantum enhanced interferometry (Kacprowicz *et al.*, 2010; Spagnolo *et al.*, 2012; Vitelli *et al.*, 2010), and most notably, when analyzing bounds on quantum enhancement in gravitational-wave detectors (Demkowicz-Dobrzański *et al.*, 2013).

Loss decoherence map  $\Lambda$  may be formally described using the following set of Kraus operators (Dorner *et al.*,

2009):

$$K_{l_a, l_b} = \sqrt{\frac{(1-\eta_a)^{l_a}}{l_a!}} \eta_a^{\hat{a}^\dagger \hat{a}} \hat{a}^{l_a} \sqrt{\frac{(1-\eta_b)^{l_b}}{l_b!}} \eta_b^{\hat{b}^\dagger \hat{b}} \hat{b}^{l_b} \quad (126)$$

where the values of index  $l_{a/b}$  corresponds to the number of photons lost in mode  $a/b$  respectively. For a general  $N$ -photon input state of the form (120), the density matrix representing the output state of the lossy interferometer reads

$$\begin{aligned} \rho_\varphi &= \Lambda(U_\varphi \rho_{\text{in}} U_\varphi^\dagger) = \\ &= \bigoplus_{N'=0}^N \sum_{\substack{l_a=0 \\ (l_b=N-N'-l_a)}}^{N-N'} p_{l_a, l_b} |\xi_{l_a, l_b}(\varphi)\rangle \langle \xi_{l_a, l_b}(\varphi)|, \end{aligned} \quad (127)$$

where  $p_{l_a, l_b} = \sum_{n=0}^N |c_n|^2 b_n^{(l_a, l_b)}$  is the binomially distributed probability of losing  $l_a$  and  $l_b$  photons in arms  $a$  and  $b$  respectively, with

$$b_n^{(l_a, l_b)} = \binom{n}{l_a} \eta_a^{n-l_a} (1-\eta_a)^{l_a} \binom{N-n}{l_b} \eta_b^{N-n-l_b} (1-\eta_b)^{l_b}, \quad (128)$$

while the corresponding conditional pure states read:

$$|\xi_{l_a, l_b}(\varphi)\rangle = \sum_{n=l_a}^{N-l_b} \frac{c_n e^{-in\varphi}}{\sqrt{p_{l_a, l_b}}} \sqrt{b_n^{(l_a, l_b)}} |n-l_a, N-n-l_b\rangle. \quad (129)$$

The direct sum in Eq. (127) indicates that the output states of different total number of surviving photons,  $N'$ , belong to orthogonal subspaces, which in principle could be distinguished by a non-demolition, photon-number counting measurement.

In the particle-approach when photons are considered as formally distinguishable particles, the loss process acts on each of the photons independently, see Fig. 9, so that the overall decoherence process has a tensor product structure  $\Lambda^{\otimes N}$ , with  $\Lambda$  being a single particle loss transformation. At the input stage, each photon occupies a two-dimensional Hilbert space spanned by vectors  $|a\rangle$ ,  $|b\rangle$  representing the photon traveling in the mode  $a/b$  respectively. In order to describe loss, however, and formally keep the number of particles constant, one needs to introduce a third photonic state at the output,  $|\text{vac}\rangle$ , representing the state of the photon being lost. Then formally,  $\Lambda$  maps states from the input two-dimensional Hilbert space to the output three-dimensional Hilbert space, and can be fully specified by means of the Kraus representation,  $\Lambda(\rho) = \sum_{i=1}^3 K_i \rho K_i^\dagger$ , where  $K_1$ ,  $K_2$ ,  $K_3$  are given respectively by the following matrices:

$$\begin{pmatrix} \sqrt{\eta_a} & 0 \\ 0 & \sqrt{\eta_b} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \sqrt{1-\eta_a} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \sqrt{1-\eta_b} \end{pmatrix}. \quad (130)$$

Intuitively, the above Kraus operators account for no photon loss, photon loss in mode  $a$  and photon loss in

mode  $b$  respectively. When applied to symmetrized input states, this loss model yields output states

$$\rho_\varphi = \Lambda^{\otimes N}(U_\varphi^{\otimes N} \rho_{\text{in}} U_\varphi^{\dagger \otimes N}) = \Lambda_\varphi^{\otimes N}(\rho_{\text{in}}), \quad (131)$$

equivalent to the ones given in Eq. (127), where  $U_\varphi$  is a single photon phase shift operation  $U_\varphi = e^{-i\varphi\hat{\sigma}_z/2}$ .

### 3. Imperfect visibility

In real-life optical interferometric experiments, it is always the case that the light beams employed do not contribute completely to the interference pattern. Due to spatiotemporal or polarization mode-mismatch, caused for example by imperfect wave-packet preparation or misalignment in the optical elements, the visibility of the interference pattern is diminished (Leonhardt, 1997; Loudon, 2000). This effect may be formally described as an effective loss of coherence between the two arms  $a$  and  $b$  of an interferometer.

Consider a single photon in a superposition state of being in modes  $a$  and  $b$  respectively:  $|\psi\rangle = \alpha|a\rangle + \beta|b\rangle$ . If other degrees of freedom such as e.g. polarization, temporal profile etc. were identical for the two states  $|a\rangle$ ,  $|b\rangle$ , we could formally write  $(\alpha|a\rangle + \beta|b\rangle)|0\rangle_X$ , where  $|0\rangle_X$  denotes the common state of additional degrees of freedom. Loss of coherence may be formally described as the transformation of the state  $|\psi\rangle|0\rangle_X$  into

$$|\Psi\rangle = \alpha \left( \sqrt{\eta}|a\rangle|0\rangle_X + \sqrt{1-\eta}|a\rangle|+\rangle_X \right) + \beta \left( \sqrt{\eta}|b\rangle|0\rangle_X + \sqrt{1-\eta}|b\rangle|-\rangle_X \right), \quad (132)$$

where  $|+\rangle_X$ ,  $|-\rangle_X$  are states orthogonal to  $|0\rangle_X$ , corresponding to photon traveling in e.g. orthogonal transversal spacial modes as depicted in Fig. 8 (iii), in which case parameter  $\eta$  can be interpreted as transmission of fictitious beam splitters that split the light into two orthogonal modes. Assuming we do not control the additional degrees of freedom the effective state of the photon is obtained by tracing out the above state over  $X$ , obtaining the effective single-photon decoherence map:

$$\Lambda(|\psi\rangle\langle\psi|) = \text{Tr}_X(|\Psi\rangle\langle\Psi|) = \begin{pmatrix} |\alpha|^2 & \alpha\beta^*\eta \\ \alpha^*\beta\eta & |\beta|^2 \end{pmatrix}, \quad (133)$$

where the off-diagonal terms responsible for coherence, are attenuated by coefficient  $\eta$ , what corresponds to the standard dephasing map (Nielsen and Chuang, 2000). Written using the Kraus representation, the above map reads:

$$\Lambda(\rho) = \sum_{i=1}^2 K_i \rho K_i^\dagger, \quad K_1 = \sqrt{\frac{1+\eta}{2}} \mathbb{1}, \quad K_2 = \sqrt{\frac{1-\eta}{2}} \sigma_z. \quad (134)$$

Note that similarly to the loss model we have modeled the noise with the use of fictitious beam-splitters to visualize the effects of decoherence. Now, as we know that a beam-splitter acts on the photons contained in its two-mode input state in an uncorrelated manner, the effective map on the full  $N$ -photon input state is  $\Lambda^{\otimes N}$ . In case of atomic systems, this would be a typical local dephasing model describing uncorrelated loss of coherence between the two relevant atomic levels (Huelga *et al.*, 1997). Still, there is a substantial difference from the loss models as the dephased photons are assumed to remain within the spatially confined beams of the interferometer arms.

We can relate the two models by a simple observation, namely that if the photons lost in the loss model with  $\eta_a = \eta_b = \eta$  were incoherently injected back into the arms of the interferometer, we would recover the local dephasing model with the corresponding parameter  $\eta$ . It should therefore come as no surprise, when we derive bounds on precisions for the two models in Sec. VI.B.1 and Sec. VI.B.2, that for the same  $\eta$  the local dephasing (imperfect visibility) model implies more stringent bounds on achievable precision than the loss model. Intuitively, it is better to get rid of the photons that lost their coherence and do not carry information about the phase, rather than to inject them back into the setup.

The structure of the output state  $\rho_\varphi$  is more complex than in the case of phase diffusion and loss models. This is because the local dephasing noise not only transforms the input state into a mixed state, but due to tracing out some degrees of freedom, the output state

$$\begin{aligned} \rho_\varphi &= \Lambda_\varphi^{\otimes N}(\rho_{\text{in}}) = \\ &= \sum_{i_1, \dots, i_N=0}^1 K_{i_1} \otimes \dots \otimes K_{i_N} U_\varphi^{\otimes N} \rho_{\text{in}} U_\varphi^{\dagger \otimes N} K_{i_1}^\dagger \otimes \dots \otimes K_{i_N}^\dagger \end{aligned} \quad (135)$$

is no longer supported on the bosonic space spanned by the fully symmetric states  $|n\rangle|N-n\rangle$ . This makes it impossible to use the mode-description in characterization of the process. Even though it is possible to write down the explicit form of the above state (Fröwis *et al.*, 2014; Jarzyna and Demkowicz-Dobrzanski, 2014) decomposing the state into  $\text{SU}(2)$  irreducible subspaces, we will not present it here for the sake of conciseness, especially that it will not be needed in derivation of the fundamental bounds.

### B. Bounds in the QFI approach

Once we have the formula for the output states  $\rho_\varphi$ , given a particular decoherence model, we may use Eq. (77) to calculate QFI,  $F_Q(\rho_\varphi)$ , which sets the limit on practically achievable precision of estimation of  $\varphi$ . In order to obtain the fundamental precision bound for a given



number  $N$  of photons used, we need to maximize the resulting  $F_Q$  over input states  $|\psi\rangle_{\text{in}}$ , which will in general be very different from the NOON states which maximize the QFI in the decoherence-free case. This is due to the fact that the NOON states are extremely susceptible to decoherence, as loss of e.g. a single photon makes the state completely useless for phase sensing. Unfortunately, for mixed states, the computation of the QFI requires in general performing the eigenvalue decomposition of  $\rho_\varphi$  and such a minimization ceases to be effective for large  $N$ . Therefore, while it is relatively easy to obtain numerical bounds on precision and the form of optimal states for moderate  $N$  (Demkowicz-Dobrzański, 2010; Dorner *et al.*, 2009; Huelga *et al.*, 1997), going to the large  $N$  regime poses a huge numerical challenge, making determination of the asymptotic bounds for  $N \rightarrow \infty$  with brute force optimization methods infeasible.

Over the past few years, elegant methods have been proposed that allow to circumvent the above mentioned difficulties and obtain explicit bounds on precision based on QFI for arbitrary  $N$ , and in particular grasp the asymptotic precision scaling (Demkowicz-Dobrzański *et al.*, 2012; Escher *et al.*, 2011; Knysh *et al.*, 2014). These methods include: the *minimization over channel purifications* method (Escher *et al.*, 2011) which is applicable in general but requires some educated guess to obtain a useful bound, as well as *classical and quantum simulation* methods (Demkowicz-Dobrzański *et al.*, 2012) which are applicable when the noise acts in an uncorrelated manner on the probes, but have an advantage of being explicit and convey some additional physical intuition on the bounds derived. Description of the newly published method (Knysh *et al.*, 2014) which is based on continuous approximation of the probe states and the calculus of variations is beyond the scope of this review.

We will present the methods by applying them directly to interferometry with each of the decoherence models introduced above. We invert the order of presentation of the bounds for the decoherence models compared with the order in Sec. VI.A, as this will allow us to discuss the methods in the order of increasing complexity. The simplest of the methods, the classical simulation, will be applied to the imperfect visibility model, while the quantum simulation will be discussed in the context of loss. Finally, minimization over channel purification method will be described in the context of the phase diffusion model, to which classical and quantum simulation methods are not applicable due to noise correlations. We should note that methods of (Escher *et al.*, 2011; Knysh *et al.*, 2014) can also be successfully applied to uncorrelated noise models. Still, classical and quantum simulation approaches are more intuitive and that is why we present derivations based on them even though the other techniques yield equivalent bounds.

In order to appreciate the significance of the derived bounds, we will always compare them with the preci-

sion achievable with a state of  $N$  uncorrelated photons as given in Eq. (97). The ratio between these quantities bounds the amount of quantum-precision enhancement that can be achieved with the help of quantum correlations present in the input state of  $N$  photons.

### 1. Imperfect visibility

The fundamental QFI bound on precision in case of imperfect visibility or equivalently the local dephasing model has been derived in (Demkowicz-Dobrzański *et al.*, 2012; Escher *et al.*, 2011; Knysh *et al.*, 2014) and reads:

$$\Delta\varphi \geq \sqrt{\frac{1-\eta^2}{\eta^2}} \frac{1}{\sqrt{N}}, \quad (136)$$

where  $\eta$  is the dephasing parameter, see Sec. VI.A.3. For the optimal uncorrelated input state,  $|\psi_{\text{in}}\rangle = [(|a\rangle + |b\rangle)/\sqrt{2}]^{\otimes N}$  we get  $\Delta\varphi = 1/\sqrt{\eta^2 N}$ , and hence the quantum precision enhancement which is the ratio of the bound on precision achievable for the optimal strategy and the precision for the uncorrelated strategy is bounded by a constant factor of  $\sqrt{1-\eta^2}$ .

*a. Classical simulation method* The derivation of the formula (136) presented below makes use of the classical simulation method (Demkowicz-Dobrzański *et al.*, 2012), which requires viewing the quantum channel representing the action of the interferometer from a geometrical perspective. The set of all physical quantum channels,  $\Lambda: \mathcal{L}(\mathcal{H}_{\text{in}}) \rightarrow \mathcal{L}(\mathcal{H}_{\text{out}})$ , that map between density matrices defined on the input/output Hilbert spaces ( $\mathcal{H}_{\text{in/out}}$ ) constitutes a convex set (Bengtsson and Życzkowski, 2006). This is to say that given any two quantum channels  $\Lambda_1, \Lambda_2$ , their convex combination  $\Lambda = p\Lambda_1 + (1-p)\Lambda_2$ ,  $0 \leq p \leq 1$  is also a legitimate quantum channel. Physically  $\Lambda$  corresponds to a quantum evolution that is equivalent to a random application of  $\Lambda_1, \Lambda_2$  transformations with probabilities  $p, 1-p$  respectively.

As derived in Sec. VI.A.3, within the imperfect visibility (local dephasing) decoherence model:  $\rho_\varphi = \Lambda_\varphi^{\otimes N}(\rho_{\text{in}})$ , and hence the relevant quantum channel, has a simple tensor structure. Consider a single-photon channel  $\Lambda_\varphi$ , which  $\varphi$ -dependence we may depict as a trajectory within the set of all single-photon quantum maps, see Fig. 10. The question of sensing the parameter  $\varphi$  can now be translated to the question of determining where on the trajectory a given quantum channel  $\Lambda_\varphi$  lies.

Consider a local *classical simulation* (CS) of a quantum channel trajectory  $\Lambda_\varphi$  in the vicinity of a given point  $\varphi_0$ ,  $\varphi = \varphi_0 + \delta\varphi$  (Demkowicz-Dobrzański *et al.*, 2012; Matsumoto, 2010),

$$\Lambda_\varphi[\varrho] = \sum_x p_\varphi(x) \Pi_x[\varrho] + O(\delta\varphi^2), \quad (137)$$

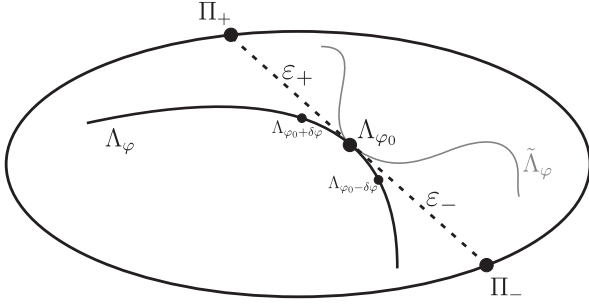


FIG. 10 The space of all quantum channels,  $\Lambda$ , which map between density matrices specified on two given Hilbert spaces,  $\Lambda : \mathcal{L}(\mathcal{H}_{\text{in}}) \rightarrow \mathcal{L}(\mathcal{H}_{\text{out}})$ , represented as a *convex* set. The estimated parameter  $\varphi$  specifies a trajectory,  $\Lambda_\varphi$  (*black curve*), in such a space. From the point of view of the QFI, any two channel trajectories, e.g.  $\Lambda_\varphi$  and  $\tilde{\Lambda}_\varphi$  (*gray curve*), are equivalent at a given  $\varphi_0$  as long as they and their first derivatives with respect to  $\varphi$  coincide there. Moreover, they can be optimally *classically simulated* at  $\varphi_0$  by mixing two channels lying on the intersection of the tangent to the trajectory and the boundary of the set:  $\Pi_\pm$ .

which represents the variation of the channel up to the first order in  $\delta\varphi$  as a classical mixture of some  $\varphi$ -independent channels  $\{\Pi_x\}_x$  where the  $\varphi$  dependence is present only in the mixing probabilities  $p_\varphi(x)$ . Under such a construction the random variable  $X$  distributed according to  $p_\varphi(x)$  specifies probabilistic choice of channels  $\Pi_x$  that reproduces the local action of  $\Lambda_\varphi$  in the vicinity of  $\varphi_0$ . Crucially, the QFI is a *local* quantity—see discussions in Sec. IV.B.1—and at a given point  $\varphi_0$  is a function only of the quantum state considered and its first derivative with respect to the estimated parameter. Consequently, when considering the parameter being encoded in a quantum channel, all the channel trajectories at a given point  $\varphi_0$  are equivalent from the point of view of QFI if they lead to density matrices that are identical up to the first order in  $\delta\varphi$ . In other words we can replace  $\Lambda_\varphi$  with any  $\tilde{\Lambda}_\varphi$  and obtain the same QFI at a given point  $\varphi_0$  provided  $\Lambda_{\varphi_0} = \tilde{\Lambda}_{\varphi_0}$  and  $\frac{d\Lambda_\varphi}{d\varphi} = \frac{d\tilde{\Lambda}_\varphi}{d\varphi}\bigg|_{\varphi=\varphi_0}$ , see Fig. 10. This means that when constructing a local CS of the quantum channel  $\Lambda_\varphi$  at  $\varphi_0$ , we need only to satisfy  $\sum_x p_{\varphi_0}(x)\Pi_x = \Lambda_{\varphi_0}$ , as well as  $\sum_x \frac{dp_\varphi(x)}{d\varphi}\bigg|_{\varphi=\varphi_0}\Pi_x = \frac{d\Lambda_\varphi}{d\varphi}\bigg|_{\varphi=\varphi_0}$ .

Crucially, as the maps  $\Lambda_\varphi$  in Eq. (135) act *independently* on each photon, we can simulate the overall  $\Lambda_\varphi^{\otimes N}$  with  $N$  *independent* random variables,  $X^N$ , associated with each channel. The estimation procedure can now be described as

$$\varphi \rightarrow X^N \rightarrow \Lambda_\varphi^{\otimes N} \rightarrow \Lambda_\varphi^{\otimes N}[\psi_{\text{in}}] \rightarrow \tilde{\varphi}, \quad (138)$$

where  $N$  classical random variables are employed to generate the desired quantum map  $\Lambda_\varphi^{\otimes N}$ . It is clear that a strategy in which we could infer the parameter directly

from  $X^N$ , i.e.  $\varphi \rightarrow X^N \rightarrow \tilde{\varphi}$ , can perform only better than the scheme where the information about  $\varphi$  is firstly encoded into the quantum channel which acts on the probe state and afterwards decoded from the measurement results performed on the output state. This way, we may always construct a classically scaling lower bound on the precision, or equivalently an upper bound on the QFI of  $\rho_\varphi$  (135):

$$F_Q[\rho_\varphi] \leq F_{\text{cl}}[p_\varphi^N] = N F_{\text{cl}}[p_\varphi], \quad (139)$$

which is determined by the classical FI (54) evaluated for the probability distribution  $p_\varphi(X)$ . Importantly, Demkowicz-Dobrzański *et al.* (2012) have shown that for the estimation problems in which the parameter is unitarily encoded, it is always optimal to choose a CS depicted in Fig. 10, which employs for each  $\varphi$  only two channels  $\Pi_\pm$  lying at the points of intersection of the tangent to the trajectory with the boundary of the quantum maps set. Such an *optimal* CS leads to the tightest upper bound specified in Eq. (139):  $F_Q[\rho_\varphi] \leq N/(\varepsilon_+ + \varepsilon_-)$ , where  $\varepsilon_\pm$  are the “distances” to the boundary marked in Fig. 10,  $\Pi_\pm = \Lambda_{\varphi_0} \pm \varepsilon_\pm \frac{d\Lambda_\varphi}{d\varphi}\bigg|_{\varphi=\varphi_0}$ .

Looking for  $\varepsilon_\pm$  parameters amounts to a search of the distances one can go along the tangent line to the trajectory of  $\Lambda_\varphi$  so that the corresponding map is still a physical quantum channel, i.e. a completely positive trace preserving map. This is easiest to do making use of the Choi-Jamiołkowski isomorphism (Choi, 1975; Jamiołkowski, 1972) which states that with each quantum channel,  $\Lambda : \mathcal{L}(\mathcal{H}_{\text{in}}) \rightarrow \mathcal{L}(\mathcal{H}_{\text{out}})$ , we can associate a positive operator  $\Omega_\Lambda \in \mathcal{L}(\mathcal{H}_{\text{out}} \otimes \mathcal{H}_{\text{in}})$ , so that  $\Omega_\Lambda = (\Lambda \otimes \mathcal{I})(|I\rangle\langle I|)$ , where  $|I\rangle = \sum_{i=1}^{\dim(\mathcal{H}_{\text{in}})} |i\rangle|i\rangle$  is a maximally entangled state on  $\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{in}}$ , while  $\mathcal{I}$  is the identity map on  $\mathcal{L}(\mathcal{H}_{\text{in}})$ . Since the complete positivity of  $\Lambda$  is equivalent to positivity of the  $\Omega_\Lambda$  operator, one needs to analyze  $\Omega_{\Lambda_{\varphi_0}} \pm \varepsilon_\pm \frac{d\Omega_\Lambda}{d\varphi}\bigg|_{\varphi=\varphi_0}$  and find maximum  $\varepsilon_\pm$  so that the above operator is still positive-semidefinite.

Taking the explicit form of the  $\Lambda_\varphi$  for the case of optical interferometry with imperfect visibility, see Sec. VI.A.3, one can show that  $\varepsilon_\pm = \sqrt{1 - \eta^2}/\eta$  (Demkowicz-Dobrzański *et al.*, 2012), which yields the ultimate quantum limit on precision given by Eq. (136).

## 2. Photonic losses

The expression for the QFI of the output state (127) in the asymptotic limit of large  $N$  has been first derived by (Knysh *et al.*, 2011). Yet, the general frameworks proposed by (Demkowicz-Dobrzański *et al.*, 2012; Escher *et al.*, 2011) for generic decoherence allowed to reconstruct this bound on precision with the following result:

$$\Delta\varphi \geq \frac{1}{2} \left( \sqrt{\frac{1 - \eta_a}{\eta_a}} + \sqrt{\frac{1 - \eta_b}{\eta_b}} \right) \frac{1}{\sqrt{N}}, \quad (140)$$

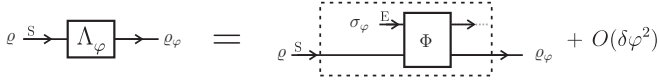


FIG. 11 The Quantum Simulation (QS) of a channel. The action of the channel  $\Lambda_\varphi$  is simulated up to the first order in the vicinity of a given point  $\varphi_0$  using a  $\varphi$ -independent  $\Phi$  channel and an auxiliary state  $\sigma_\varphi$  that contains the *full* information about the estimated parameter  $\varphi$ .

where  $\eta_a, \eta_b$  are transmission in the two arms of the interferometer respectively, see VI.A.2. This bound simplifies to

$$\Delta\varphi \geq \sqrt{\frac{1-\eta}{\eta N}} \quad (141)$$

in the case of equal losses, and since the precision achievable with uncorrelated states is given by  $1/\sqrt{\eta N}$ , the maximal quantum-enhancement factor is  $\sqrt{1-\eta}$ . In the following, we derive the above bounds using the quantum simulation approach of (Demkowicz-Dobrzański *et al.*, 2012; Kołodyński and Demkowicz-Dobrzański, 2013).

*a. Quantum Simulation method* Unfortunately, in the case of loss the simple CS method yields a trivial bound  $\Delta\varphi \geq 0$ , since the tangent distances to the boundary of the set of quantum channels are  $\varepsilon_\pm = 0$  in this case. It is possible, however, to derive a useful bound via the *Quantum Simulation (QS) method* which is a natural generalization of the CS method. The QS method has been described in detail and developed for general metrological schemes with uncorrelated noise by Kołodyński and Demkowicz-Dobrzański (2013) stemming from the works of Demkowicz-Dobrzański *et al.* (2012) and Matsumoto (2010).

As shown in Fig. 11, local QS amounts to re-expressing the action of  $\Lambda_\varphi$  for  $\varphi = \varphi_0 + \delta\varphi$  by a larger  $\varphi$ -independent map  $\Phi$  that also acts on the auxiliary  $\varphi$ -dependent input  $\sigma_\varphi$ , up to the first order in  $\delta\varphi$ :

$$\Lambda_\varphi[\varrho] = \text{Tr}_E \Phi[\sigma_\varphi \otimes \varrho] + O(\delta\varphi^2). \quad (142)$$

Note that for  $\sigma_\varphi = \sum_x p_\varphi(x) |x\rangle\langle x|$ , and  $\Phi = |x\rangle\langle x| \otimes \Pi_x$ , QS becomes equivalent to the CS of Eq. (137), so that CS is indeed a specific instance of the more general QS. An analogous reasoning as in the case of CS leads to the conclusion that we may upper-bound the QFI of the overall output state, here (127) for the case of losses, as

$$\begin{aligned} F_Q[\Lambda_\varphi^{\otimes N}[\psi]_{\text{in}}\langle\psi|]] &= F_Q[\text{Tr}_E\{\Phi^{\otimes N}[\sigma_\varphi^{\otimes N} \otimes |\psi\rangle_{\text{in}}\langle\psi|]\}] \\ &\leq F_Q[\sigma_\varphi^{\otimes N}] = N F_Q[\sigma_\varphi], \end{aligned}$$

since  $\text{Tr}_E \Phi^{\otimes N}[\cdot]$  is just a parameter independent map, under which the overall QFI may only decrease—see

Eq. (82). Last equality follows from the additivity property of the QFI, which, similarly to Eq. (139), constrains  $F_Q[\rho_\varphi^N]$  to scale at most linearly for large  $N$ . Similarly to the case of CS, in order to get the tightest bound one should find QS that yields the smallest  $F_Q[\sigma_\varphi]$ , which in principle is a non trivial task.

Fortunately, Kołodyński and Demkowicz-Dobrzański (2013) have demonstrated that the search for the optimal channel QS corresponds to the optimization over the Kraus representation of a given channel. Without loss of generality we may assume that  $\sigma_\varphi = |\varphi\rangle\langle\varphi|$  is a pure  $\varphi$ -dependent state while  $\Phi[\cdot] = \mathcal{U} \cdot \mathcal{U}^\dagger$  is unitary. For a given QS we may write the corresponding Kraus representation of the channel by choosing a particular basis  $|i\rangle_E$  in the E space:  $K_i(\varphi) = {}_E\langle i|\mathcal{U}|\varphi\rangle_E$ . In order for the QS to be valid, these Kraus operators should correspond to a legitimate Kraus representation of the channel  $\Lambda_\varphi[\cdot] = \sum_i K_i(\varphi) \cdot K_i(\varphi)^\dagger$ . Two Kraus representation of a given quantum channel are equivalent if and only if they are related by a unitary matrix  $u$ :

$$\tilde{K}_i(\varphi) = \sum_j u_{ij}(\varphi) K_j(\varphi), \quad (143)$$

which may in principle be also  $\varphi$  dependent. Since we require QS to be only locally valid in the vicinity of  $\varphi_0$ , the above equation as well as its first derivative needs to be fulfilled only at  $\varphi_0$ . Because of that, the search for the optimal Kraus representation  $\tilde{K}_i$  (or equivalently the optimal QS) may be restricted to the class of transformations where  $u(\varphi) = e^{i(\varphi-\varphi_0)\mathbf{h}}$  with  $\mathbf{h}$  being any Hermitian matrix that shifts the relevant derivative of  $K_i(\varphi)$  at  $\varphi_0$ , so that  $\tilde{K}_i(\varphi_0) = K_i(\varphi_0)$  and  $\dot{\tilde{K}}_i(\varphi_0) = \dot{K}_i(\varphi_0) + i \sum_j \mathbf{h}_{ij} K_j(\varphi_0)$ . As shown by (Kołodyński and Demkowicz-Dobrzański, 2013), the problem of finding the optimal QS i.e. the minimal  $F_Q[|\varphi\rangle\langle\varphi|]$  which we term as the  $F_{\text{QS}}$ , can be formally rewritten as

$$\begin{aligned} F_{\text{QS}} &= \min_{\mathbf{h}} s \quad \text{s.t.} \\ \sum_i \dot{\tilde{K}}_i(\varphi_0)^\dagger \dot{\tilde{K}}_i(\varphi_0) &= \frac{s}{4} \mathbb{1}_2, \quad \sum_i \dot{\tilde{K}}_i(\varphi_0)^\dagger \tilde{K}_i(\varphi_0) = \mathbf{0}, \end{aligned} \quad (144)$$

where the parameter  $s$  has the interpretation of  $F_Q[|\varphi\rangle\langle\varphi|]$  for the particular QS at  $\varphi_0$  and the constraints imposed in Eq. (144) are necessary and sufficient for the QS required transformation  $\mathcal{U}$  and the state  $|\varphi\rangle$  to exist.

The above optimization problem may not always be easy to solve. Still, its relaxed version:

$$\min_{\mathbf{h}} \left\| \sum_i \dot{\tilde{K}}_i(\varphi_0)^\dagger \dot{\tilde{K}}_i(\varphi_0) \right\|, \quad \sum_i \dot{\tilde{K}}_i(\varphi_0)^\dagger \tilde{K}_i(\varphi_0) = \mathbf{0}, \quad (145)$$

where  $\|\cdot\|$  is the operator norm, can always be cast in the form of an explicit semi-definite program, which can be easily solved numerically (Demkowicz-Dobrzański *et al.*, 2012). Numerical solution of the semi-definite program

provides a form of the optimal  $h$  which may then be taken as an ansatz for further analytical optimization.

Plugging in the Kraus operators  $K_i(\varphi) = K_i U_\varphi$  representing the lossy interferometer, see Eq. (130), and following the above described procedure one obtains

$$F_{\text{QS}} = \frac{4}{\left(\sqrt{\frac{1-\eta_a}{\eta_a}} + \sqrt{\frac{1-\eta_b}{\eta_b}}\right)^2} \quad (146)$$

for the optimal  $h$  given by

$$h_{\text{opt}} = -\frac{1}{8} \text{diag} \left\{ \chi, \frac{\eta_a}{1-\eta_a} \left( \frac{4}{\eta_a} - \chi \right), -\frac{\eta_b}{1-\eta_b} \left( \frac{4}{\eta_b} + \chi \right) \right\}, \quad (147)$$

where  $\chi = F_{\text{QS}} \frac{\eta_b - \eta_a}{\eta_a \eta_b}$ . This indeed reproduces the bounds given in (140).

In order to provide the reader with a simple intuition concerning the QS method, we shall present an elementary construction of the QS for lossy interferometer in the special case of  $\eta_a = \eta_b = 1/2$ . In this case the bound (140) yields  $\Delta\varphi \geq 1/\sqrt{N}$  which implies that using optimal entangled probe state at the input under 50% losses cannot beat the precision which can be obtained by uncorrelated probes in ideal scenario of no losses.

Consider the action of the single photon lossy channel  $\Lambda_\varphi$  on the pure input state  $|\psi\rangle = \alpha|a\rangle + \beta|b\rangle$ :

$$\Lambda_\varphi(|\psi\rangle\langle\psi|) = \frac{1}{2}|\psi_\varphi\rangle\langle\psi_\varphi| + \frac{1}{2}|\text{vac}\rangle\langle\text{vac}| \quad (148)$$

with  $|\psi_\varphi\rangle = \alpha e^{i\varphi}|a\rangle + \beta|b\rangle$ , which represents 1/2 probability of photon sensing the phase undisturbed and the 1/2 probability of the photon being lost. Let the auxiliary state for QS be  $|\varphi\rangle = (e^{i\varphi}|0\rangle + |1\rangle)/\sqrt{2}$ . The joined input + auxiliary state reads:

$$|\phi\rangle|\psi\rangle = \frac{1}{\sqrt{2}} (\alpha e^{i\varphi}|0\rangle|a\rangle + \beta|1\rangle|b\rangle) + \frac{1}{\sqrt{2}} (\alpha e^{i\varphi}|0\rangle|b\rangle + \beta|1\rangle|a\rangle). \quad (149)$$

The map  $\Phi$  realizing the QS consists now of two steps. First the controlled NOT operation is performed with the auxiliary system being the target qubit, this transform the above state to:  $\frac{1}{\sqrt{2}}|0\rangle(\alpha e^{i\varphi}|a\rangle + \beta|b\rangle) + \frac{1}{\sqrt{2}}|1\rangle(\alpha e^{i\varphi}|b\rangle + \beta|a\rangle)$ . The second step is the measurement of the auxiliary system. If the result  $|0\rangle$  is measured (probability 1/2), the system is left in the correct state  $|\psi_\varphi\rangle$  and the map leaves it unchanged, if the  $|1\rangle$  is measured the state of the photon is not the desired one, in which case the map returns the  $|\text{vac}\rangle$  state. This map is therefore a proper QS of the desired lossy interferometer transformation for  $\eta_a = \eta_b = 1/2$ . Since the auxiliary state employed in this construction was  $|\varphi\rangle = (e^{i\varphi}|0\rangle + |1\rangle)/\sqrt{2}$  for which  $F_Q(|\varphi\rangle\langle\varphi|^{\otimes N}) = N$  this leads to the anticipated result  $\Delta\varphi \geq 1/\sqrt{N}$ .

### 3. Phase diffusion

Since the phase diffusion model, see Sec. VI.A.1, is an example of a correlated noise model, it cannot be approached with the CS and QS methods. The study of the behavior of the QFI within the phase-diffusion model was for the first time carried out by (Genoni *et al.*, 2011) considering indefinite-photon-number Gaussian input states and studied numerically the achievable precision and the structure of optimal input states. Yet, the fundamental analytical bounds on precision have not been verified until Escher *et al.* (2012), where the phase noise has been approached using the minimization over purifications method of Escher *et al.* (2011) and most recently using the calculus of variations approach of Knysh *et al.* (2014).

*a. Minimization over purification method* The minimization over purification method of Escher *et al.* (2011) is based on the observation, already mentioned in Sec. IV.B.1, that QFI for a given mixed quantum state, here  $\rho_\varphi$  (127), is not only upper bounded by the QFI of any of its purifications, but there always exists an optimal purification,  $|\Psi_\varphi^{\text{opt}}\rangle$ , for which  $F_Q[\rho_\varphi] = F_Q[|\Psi_\varphi^{\text{opt}}\rangle]$ , where  $\rho_\varphi = \text{Tr}_E\{|\Psi_\varphi^{\text{opt}}\rangle\langle\Psi_\varphi^{\text{opt}}|\}$ . As such statement does not rely on the form of the transformation  $|\psi\rangle_{\text{in}} \rightarrow \rho_\varphi$  but rather on the properties of the output state itself, the framework of Escher *et al.* (2011) in principle does not put any constraints on the noise-model considered. Note that, even if the optimal purification itself is difficult to find, any suboptimal purification yields a legitimate upper bound on the QFI and hence may provide a non-trivial precision bound.

In order to get a physical intuition regarding the purification method, consider a physical model of the phase diffusion where light is being reflected from a mirror which position fluctuations are randomly changing the effective optical length. Formally, the model amounts to coupling the phase delay generator  $\hat{J}_z$  to the mirror position quadrature  $\hat{x}_E = \frac{1}{\sqrt{2}}(\hat{a}_E + \hat{a}_E^\dagger)$  (Escher *et al.*, 2012). Assuming the mirror, serving as the environment E, to reside in the ground state of a quantum oscillator  $|0\rangle_E$  before interaction with the light beam, the pure output state reads:

$$|\Psi_\varphi\rangle = e^{-i\varphi\hat{J}_z} e^{i\sqrt{2}\Gamma\hat{J}_z\hat{x}_E} |\psi\rangle_{\text{in}} |0\rangle_E. \quad (150)$$

Thanks to the fact that  $|\langle x|0\rangle|^2 = e^{-x^2}/\sqrt{\pi}$ , the reduced state

$$\rho_\varphi = \int_{-\infty}^{\infty} dx \, {}_E\langle x|\Psi_\varphi\rangle\langle\Psi_\varphi|x\rangle_E \quad (151)$$

indeed coincides with the correct output phase diffused state (125). Therefore, this is a legitimate purification



of the interferometer output state in presence of phase diffusion.

Consider now another purification  $|\tilde{\Psi}_\varphi\rangle = e^{i\delta\varphi\hat{H}_E}|\Psi_\varphi\rangle$  generated by a local ( $\varphi = \varphi_0 + \delta\varphi$ ) rotation of the mirror modes, i.e. a unitary operation on the system E. We look for a transformation of the above form which hopefully erases as much information on the estimated phase as possible, so that QFI for the purified state  $|\tilde{\Psi}_\varphi\rangle$  will be minimized leading to the best bound on the QFI of  $\rho_\varphi$ . Choosing  $\hat{H}_E = \lambda\hat{p}_E$  we obtain the following upper bound on the QFI

$$F_Q[\rho_\varphi] \leq \min_\lambda \{F_Q[e^{i\varphi\lambda\hat{p}_E}|\Psi(\varphi)\rangle]\} = \\ = \min_\lambda \left\{ 2\lambda^2 + 4 \left(1 - \sqrt{2\Gamma\lambda}\right)^2 \Delta^2 J_z \right\} = \frac{4\Delta^2 J_z}{1 + 4\Gamma\Delta^2 J_z},$$

and thus a lower limit on the precision

$$\Delta\varphi \geq \sqrt{\Gamma + \frac{1}{4\Delta^2 J_z}} = \sqrt{\Gamma + \frac{1}{N^2}}, \quad (152)$$

where we plugged in  $\Delta^2 J_z = N^2/4$  corresponding to the NOON state which maximized the variance for  $N$ -photon states. See also an alternative derivation of the above result that has been proposed recently in (Macieszczak, 2014). Crucially, the above result proves that the phase diffusion constrains the error to approach a constant value  $\sqrt{\Gamma}$  as  $N \rightarrow \infty$ , which does not vanish in the asymptotic limit, what contrasts the  $1/\sqrt{N}$  behavior characteristic for uncorrelated noise models. Note also that due to the correlated character of the noise, the bound (152) predicts that it may be more beneficial to perform the estimation procedure on a group of  $k$  particles and then repeat the procedure independently  $\nu$  times obtaining  $1/\sqrt{\nu}$  reduction in estimation error, rather than employing  $N = k\nu$  in a single experimental shot (Knysh *et al.*, 2014).

Only very recently, the exact ultimate quantum limit for the  $N$ -photon input states has been derived in (Knysh *et al.*, 2014)

$$\Delta\varphi \geq \sqrt{\Gamma + \frac{\pi^2}{N^2}}, \quad (153)$$

showing that the previous bound was not tight, with the second term following the HL-like asymptotic scaling of the noiseless decoherence-free Bayesian scenario stated in Eq. (107). In fact, as proven by (Knysh *et al.*, 2014), the optimal states of the noiseless Bayesian scenario, i.e. the sine states (105), attain the above correct quantum limit. In Sec. VI.C.2, we show that within the Bayesian approach, with the phase-diffusion effects incorporated, the sine states are always the optimal inputs.

### C. Bayesian approach

Minimizing the average Bayesian cost, as given by Eq. (93), over input probe states  $|\psi\rangle_{\text{in}}$  is in general more demanding than minimization of the QFI due to the fact that it is not sufficient to work within the local regime and analyze only the action of a channel and its first derivative at a given estimation point as in the QFI approach. For this reason we do not apply the Bayesian approach to the imperfect visibility model as obtaining the bounds requires a significant numerical and analytical effort (Jarzyna and Demkowicz-Dobrzański, 2014), and constrain ourselves to loss and phase diffusion models.

#### 1. Photonic losses

The optimal Bayesian performance of  $N$ -photon states has been studied by Kołodyński and Demkowicz-Dobrzański (2010). Assuming the natural cost function (67) and the flat prior phase distribution, the average cost (93) reads:

$$\langle C \rangle = \text{Tr}\{\langle \rho_\varphi \rangle_C \Xi\} \quad (154)$$

where  $\langle \rho_\varphi \rangle_C = 4 \int \frac{d\varphi}{2\pi} \rho_\varphi \sin^2(\frac{\varphi}{2})$  and  $\rho_\varphi$  is given by (127). The optimal measurement seed operator  $\Xi$  can be found analogously as in the decoherence-free case. The block diagonal form of  $\rho_\varphi$ , implies that without losing optimality one can assume  $\Xi = \bigoplus_{N'=0}^N |e_{N'}\rangle\langle e_{N'}|$  with  $|e_{N'}\rangle = \sum_{n=0}^{N'} |n, N'-n\rangle$ . Physically, the block-diagonal structure of  $\Xi$  indicates that the optimal covariant measurement requires a non-demolition photon number measurement to be performed before carrying out any phase measurements, so that the orthogonal subspaces, labeled by the number of surviving photons  $N'$ , may be firstly distinguished, and subsequently the measurement which is optimal in the lossless case is performed (Kołodyński and Demkowicz-Dobrzański, 2010). Plugging in the explicit form of  $\Xi$  together with the explicit form of the output state  $\rho_\varphi$ , we arrive at

$$\langle C \rangle = 2 - \mathbf{c}^T \mathbf{A} \mathbf{c}, \quad A_{n,n-1} = A_{n-1,n} = \sum_{l_a, l_b=0}^{n, N-n} \sqrt{b_n^{(l_a, l_b)} b_{n-1}^{(l_a, l_b)}}, \quad (155)$$

where  $\mathbf{A}$  is a symmetric  $(N+1) \times (N+1)$  matrix that is non-zero only on its first off-diagonals,  $b_n^{(l_a, l_b)}$  are the binomial coefficients previously defined in Eq. (128), while  $\mathbf{c}$  is a state vector containing coefficients  $c_n$  of the  $N$ -photon input state (120).

The *minimal* average cost (154) for the lossy interferometer then equals  $\langle C \rangle_{\min} = 2 - \lambda_{\max}$ , where  $\lambda_{\max}$  is the maximal eigenvalue of the matrix  $\mathbf{A}$  and the corresponding eigenvector  $\mathbf{c}_{\max}$  provides the optimal input

state coefficients.  $\langle C \rangle_{\min}$  quantifies the maximal achievable precision and in the  $N \rightarrow \infty$  limit may be interpreted as the average MSE (60) due to the convergence of the cost function (67) to the squared distance as  $\tilde{\varphi} \rightarrow \varphi$ .

The procedure described above allows only to obtain numerical values of the achievable precision, and ceases to be feasible for  $N \rightarrow \infty$ . The main result of (Kołodziej and Demkowicz-Dobrzański, 2010) was to construct a valid analytical lower bound on the minimal average cost (154):

$$\langle C \rangle_{\min} \geq 2 \left[ 1 - A_{\max} \cos \left( \frac{\pi}{N+2} \right) \right], \quad (156)$$

where  $A_{\max} = \max_{1 \leq n \leq N} \{A_{n,n-1}\}$  is the largest element of the matrix  $\mathbf{A}$ , contained within its off-diagonal entries (155). The bound yields exactly the same formula as the QFI bound (140), proving that in this case the Bayesian and QFI approaches are equivalent:

$$\Delta\varphi \approx \sqrt{\langle C \rangle} \geq \sqrt{\langle C \rangle_{\min}} \geq \frac{1}{2} \left( \sqrt{\frac{1-\eta_a}{\eta_a}} + \sqrt{\frac{1-\eta_b}{\eta_b}} \right) \frac{1}{\sqrt{N}}, \quad (157)$$

where  $\approx$  represents the fact that Bayesian cost approximated the variance only in the limit of large  $N$ . The fact that both approaches lead to the same ultimate bounds on precision suggests that the optimal input states may be approximated for  $N \rightarrow \infty$  up to an arbitrary good precision with states manifesting only local finite-number of particle correlations and may in particular be efficiently simulated with the concept of matrix-product states (Jarzyna and Demkowicz-Dobrzański, 2013; Jarzyna and Demkowicz-Dobrzański, 2014).

## 2. Phase diffusion

Similarly to the case of losses discussed in the previous section, we study the estimation precision achieved within the Bayesian approach but in the presence of *phase diffusion*. The analysis follows exactly in the same way, so that likewise assuming *no prior knowledge* and the natural cost function introduced in Eq. (93) the formula for the average cost reads

$$\langle C \rangle = \text{Tr} \{ \langle \rho_\varphi \rangle_C \Xi \} = 2 - \mathbf{c}^T \mathbf{A} \mathbf{c}, \quad (158)$$

where this time one may think of the effective state, as of the input state  $\rho_{\text{in}}$  which is firstly averaged over the Gaussian distribution dictated by the evolution (124) and then over the cost function in accordance with Eq. (93). The optimal seed element of the covariant POVM is identical as in the decoherence-free case  $\Xi = |e_N\rangle\langle e_N|$  and the matrix  $\mathbf{A}$  possesses again only non-zero entries on its first off-diagonals, but this time all of them are equal to  $e^{-\frac{\Gamma}{2}}$ . As a result, the *minimal* average cost (158) may be evaluated analytically following exactly the calculation

of (Berry and Wiseman, 2000) for the noiseless scenario, which leads then to  $\lambda_{\max} = 2e^{-\frac{\Gamma}{2}} \cos \left( \frac{\pi}{N+2} \right)$  and hence

$$\langle C \rangle_{\min} = 2 \left[ 1 - e^{-\frac{\Gamma}{2}} \cos \left( \frac{\pi}{N+2} \right) \right] \stackrel{N \rightarrow \infty}{\approx} 2 \left( 1 - e^{-\frac{\Gamma}{2}} \right) + e^{-\frac{\Gamma}{2}} \frac{\pi^2}{N^2}. \quad (159)$$

The optimal input states are the same as in the decoherence-free case, i.e. they are the  $N$ -photon sine states of Eq. (105). Note that in contrast to the photonic loss which is an example of an uncorrelated noise, the minimal average cost (159) does not asymptotically coincide with the QFI-based precision limit (153) unless  $\Gamma \ll 1$ .

## D. Practical schemes saturating the bounds

Deriving the fundamental bounds on quantum enhanced precision in presence of decoherence is interesting in itself from a theoretical a point of view. Still, a practical question remains whether the bounds derived are saturable in practice. Note that NOON states and the sine states that are optimal in case of QFI and Bayesian approaches in the decoherence-free case are notoriously hard to prepare apart from regime of very small  $N$ . For large photon numbers, the only practically accessible states of light are squeezed Gaussian states and one of the most popular strategies in performing quantum-enhanced interferometry amounts to mixing a coherent beam with a squeezed vacuum state on the input beam splitter of the Mach-Zehnder interferometer, see Sec. III.D. We demonstrate below that in presence of uncorrelated decoherence, such as loss or imperfect visibility, this strategy is indeed optimal in the asymptotic regime of large  $N$  and allows to saturate the fundamental bounds derived above. We will not discuss the phase-diffusion noise, since the estimation uncertainty is finite in the asymptotic limit, and the issue of saturating the asymptotic bound becomes trivial as practically all states lead to the same asymptotic precision value, while saturating the bound for finite  $N$  requires the use of experimentally inaccessible sine states.

### 1. Bounds for indefinite photon number states

Derivation of the bounds presented in this section both in the QFI and Bayesian approaches assumed *definite*-photon number states at the input. We have already discussed the issue of translating the bounds from a definite photon number input state case to a general indefinite-photon number state case in Sec. V.C in the case of decoherence-free metrology, where we have observed that

due to quadratic dependence of QFI on number of photons used, maximization of QFI over states with fixed *averaged* photon number  $\langle N \rangle$  is ill defined and arbitrary high QFI are in principle achievable. Controversies related to this observation, discussed in Sec. V.C.2, are fortunately not present in the noisy metrology scenario.

For the decoherence models, analyzed in this paper, the QFI scales at most linearly with  $N$ . Following the reasoning presented in Sec. V.C, consider a mixture of different photon number states  $\sum_N p_N \rho_N$ . Since in the presence of decoherence  $F_Q(\rho_N) \leq cN$ , where  $c$  is a constant coefficient that depends on the type and strength of the noise considered, thanks to the convexity of the QFI we can write:

$$F_Q \left( \sum_N p_N \rho_N \right) \leq \sum_N p_N F_Q(\rho_N) \leq \sum_N p_N cN = c\langle N \rangle. \quad (160)$$

Hence the bounds on precision derived in Sec. VI.B.2 (losses) and Sec. VI.B.1 (imperfect visibility) are valid also under replacement of  $N$  by  $\langle N \rangle$ . Still, one may come across claims of precisions going beyond the above mentioned bounds typically by a factor of two (Aspachs *et al.*, 2009; Joo *et al.*, 2011). This is only possible, however, if classical reference beam required to perform e.g. the homodyne detection is not treated as a resource. As discussed in detail in Sec. V.C.1, we take the position that such reference beams should be treated in the same way as the light traveling through the interferometer and as such also counted as a resource.

## 2. Coherent + squeezed vacuum strategy

In section Sec. III.A, we have derived an error-propagation formula for the phase-estimation uncertainty (33) for the standard Mach-Zehnder interferometry in absence of decoherence. For this purpose we have adopted the Heisenberg picture and expressed the precision in terms of expectation values, variances and covariances of the respective angular momentum observables calculated for the input state. Here, we follow the same procedure but take additionally into account the effect of imperfect visibility (local dephasing) and loss. For simplicity, in the case of loss we restrict ourselves to equal losses in both arms. The Heisenberg picture transformation of an observable  $\hat{O}$  corresponding to a general map  $\Lambda_\varphi$  (123) reads

$$\sum_i U_\varphi^\dagger K_i^\dagger \hat{O} K_i U_\varphi = \Lambda_\varphi^*(\hat{O}), \quad (161)$$

where  $\Lambda^*$  is called the conjugated map.

For a more direct comparison with the decoherence-free formulas of Sec. III.A, we explicitly include the action of the Mach-Zehnder input and output balanced beam

splitters in the description of the state transformation—in terms of Fig. 8 this corresponds to moving  $|\psi\rangle_{\text{in}}$  to the left and  $\rho_\varphi$  to the right of the figure. In case of loss the decoherence map has the same form as given in Sec. VI.A.2, but with  $U_\varphi = e^{-i\frac{\varphi}{2}\sigma_y}$ , while in the case of imperfect visibility the Kraus operators (134) will be modified to  $K_1 = \sqrt{\frac{1+\eta}{2}}\mathbb{1}$ ,  $K_2 = \sqrt{\frac{1-\eta}{2}}\sigma_y$ , so that the local dephasing is defined with respect to the  $y$  rather than the  $z$  axis. For the two decoherence models, the resulting Heisenberg picture transformation of the  $J_z$  observable yields (Ma *et al.*, 2011):

$$\begin{aligned} \frac{\langle \hat{J}_z \rangle}{\eta} &= \cos \varphi \langle \hat{J}_z \rangle_{\text{in}} - \sin \varphi \langle \hat{J}_x \rangle_{\text{in}}, \\ \frac{\Delta^2 J_z}{\eta^2} &= f(\eta) \frac{\langle N \rangle}{4} + \cos^2 \varphi \Delta^2 J_z|_{\text{in}} + \sin^2 \varphi \Delta^2 J_x|_{\text{in}} + \\ &\quad - 2 \sin \varphi \cos \varphi \text{cov}(J_x, J_z)|_{\text{in}}. \end{aligned} \quad (162)$$

where  $f(\eta) = (1 - \eta)/\eta$  for the loss model and  $f(\eta) = (1 - \eta^2)/\eta^2$  in the case of local dephasing model. The above expressions have a clear intuitive interpretation. The signal  $\langle \hat{J}_z \rangle$  is rescaled by a factor  $\eta$  compared with the decoherence-free case, while the variance apart from the analogous rescaling is enlarged by an additional noise contribution  $f(\eta)\langle N \rangle/4$  due to lost or dephased photons.

In order to calculate the precision achievable with coherent+squeezed-vacuum strategy, we may use the already obtained quantities presented in Eq. (37). After substituting the input variances and averages into Eq. (162) and optimally setting  $\alpha = \text{Re}(\alpha)$  as before, we arrive at a modified version of the formula (38) for the phase estimation precision:

$$\begin{aligned} \Delta\varphi^{|\alpha\rangle|r} &= \\ &= \frac{\sqrt{\cot^2 \varphi (|\alpha|^2 + \frac{1}{2} \sinh^2 2r) + |\alpha|^2 e^{-2r} + \sinh^2 r + f(\eta) \frac{|\alpha|^2 + \sinh^2 r}{\sin^2 \varphi}}}{||\alpha|^2 - \sinh^2 r|}. \end{aligned} \quad (163)$$

The optimal operation points are again  $\varphi = \pi/2, 3\pi/2$ . Considering the asymptotic limit  $\langle N \rangle = |\alpha|^2 + \sinh^2 r \rightarrow \infty$  and assuming the coherent beam to carry the dominant part of the energy  $|\alpha| \gg \sinh^2 r$ , the formula for precision at the optimal operation point reads:

$$\Delta\varphi^{|\alpha\rangle|r} \approx \frac{\sqrt{\langle N \rangle e^{-2r} + f(\eta) \langle N \rangle}}{\langle N \rangle} = \frac{\sqrt{e^{-2r} + f(\eta)}}{\sqrt{\langle N \rangle}}. \quad (164)$$

Clearly, even for relatively small squeezing strength  $r$  the  $e^{-2r}$  term becomes negligible, and hence we can effectively approach arbitrary close precision given by:

$$\Delta\varphi^{|\alpha\rangle|r} \approx \frac{\sqrt{f(\eta)}}{\sqrt{\langle N \rangle}}, \quad (165)$$

which recalling the definition of  $f(\eta)$  for the two decoherence models considered coincides exactly with the

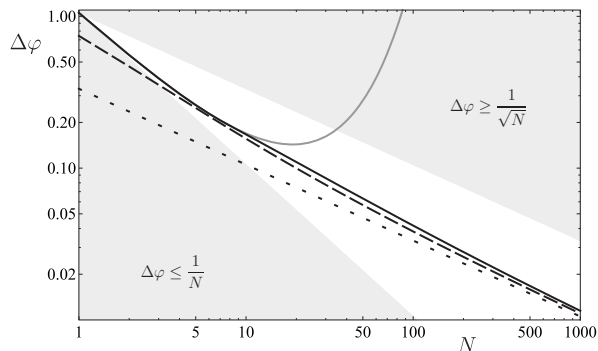


FIG. 12 The phase estimation precision of an interferometer with equal losses in both arms ( $\eta = 0.9$ ). The performance of the optimal  $N$ -photon input states (120) is shown (solid black) that indeed saturate the asymptotic quantum limit (140) (dotted):  $\sqrt{(1-\eta)/(\eta N)}$ . The NOON states (solid grey) achieve nearly optimal precision only for low  $N$  ( $\leq 10$ ) and rapidly diverge becoming out-performed by classical strategies. For comparison, the precision attained for an indefinite photon number scheme is presented, i.e. a coherent state and squeezed vacuum optimally mixed on a beam-splitter (Caves, 1981) (dashed), which in the presence of loss also saturates the asymptotic quantum limit (140).

fundamental bounds (136), (141) derived before. This proves that the fundamental bounds can be asymptotically saturated with a practical interferometric scheme. One should note that this contrasts the noiseless case and the suboptimal performance of simple estimation scheme based on the photon-number difference measurements, see Eq. (39).

To summarize the results obtained in this section, in Fig. 12, we present a plot of the maximal achievable precision for the lossy interferometer in the equal-losses scenario with  $\eta = 0.9$ , i.e.  $\Delta\varphi = 1/\sqrt{F_Q[\rho_\varphi]}$  as a function of  $N$  compared with the NOON state-based strategy as well as the asymptotic bound (140). On the one hand, the NOON states remain optimal for relatively small  $N$  ( $\leq 10$ ), for which the effects of losses may be disregarded. This fact supports the choice of NOON-like states in the quantum-enhanced experiments with small number of particles (Krischek *et al.*, 2011; Mitchell *et al.*, 2004; Nagata *et al.*, 2007; Okamoto *et al.*, 2008; Resch *et al.*, 2007; Xiang *et al.*, 2010). However, one should note that in the presence of even infinitesimal losses, the precision achieved by the NOON states quickly diverges with  $N$ , because their corresponding output state QFI,  $F_Q^{\text{NOON}} = \eta^N N^2$ , decays exponentially for any  $\eta < 1$ .

Most importantly, it should be stressed that the coherent+squeezed vacuum strategy discussed above has been implemented in recent gravitational-wave interferometry experiments (LIGO Collaboration, 2011, 2013). The main factor limiting the quantum enhancement of precision in this experiments is loss, which taking into

account detection efficiency, optical instruments imperfections and imperfect coupling was estimated at the level of 38% (LIGO Collaboration, 2011). In (Demkowicz-Dobrzański *et al.*, 2013) it has been demonstrated that the sensing precision achieved in (LIGO Collaboration, 2011) using the 10dB squeezed vacuum (corresponding to the squeezing factor  $e^{-2r} \approx 0.1$ ), was strikingly close to the fundamental bound, and only 8% further reduction in estimation uncertainty would be possible if more advanced input states of light were used.

## VII. CONCLUSIONS

In this review we have showed how the tools of quantum estimation theory can be applied in order to derive fundamental bounds on achievable precision in quantum-enhanced optical interferometric experiments. The main message to be conveyed is the fact that while the power of quantum enhancement is seriously reduced by the presence of decoherence, and in general the Heisenberg scaling cannot be reached, non-classical states of light offer a noticeable improvement in interferometric precision and simple experimental schemes may approach arbitrarily close the fundamental quantum bounds. It is also worth noting that in the presence of uncorrelated decoherence the Bayesian approaches coincide asymptotically with the QFI approaches easing the tension between this two often competing ways of statistical analysis.

We would also like to mention an inspiring alternative approach to the derivation of limits on precision of phase estimation, where the results are derived making use of information theoretic concepts such as rate-distortion theory (Nair, 2012) or entropic uncertainty relations (Hall and Wiseman, 2012). Even though the bounds derived in this way are weaker than the bounds presented in this review and obtained via Bayesian or QFI approaches, they carry a conceptual appeal encouraging to look for deeper connections between quantum estimation and communication theories.

Let us also point out, that while we have focused our discussion on optical interferometry using the paradigmatic Mach-Zehnder model, the same methods can be applied to address the problems of fundamental precision bounds in atomic interferometry (Cronin *et al.*, 2009), magnetometry (Budker and Romalis, 2007), frequency stabilization in atomic clocks (Diddams *et al.*, 2004) as well as the limits on resolution of quantum enhanced lithographic protocols (Boto *et al.*, 2000). All these setups can be cast into a common mathematical framework, see Sec. III.F, but the resulting bounds will depend strongly on the nature of dominant decoherence effects and the relevant resource limitations such as: total experimental time, light power, number of atoms etc., as well as on the chosen figure of merit. In particular, it is not excluded that in some atomic metrological scenarios



one may still obtain a better than  $1/\sqrt{N}$  of precision if decoherence is of a special form allowing for use of the decoherence-free subspaces (Dorner, 2012; Jeske *et al.*, 2013) or when its impact may be significantly reduced by considering short evolution times, in which the SQL-like scaling bounds may in principle be circumvented by: adjusting decoherence geometry (Arrad *et al.*, 2014; Chaves *et al.*, 2013; Dür *et al.*, 2014; Kessler *et al.*, 2013) or by considering Non-Markovian short-time behaviour (Chin *et al.*, 2012; Matsuzaki *et al.*, 2011).

We should also note that application of the tools presented in this review to a proper analysis of fundamental limits to the operation of quantum enhanced atomic clocks (Andrè *et al.*, 2004; Leibfried *et al.*, 2004) is not that direct as it requires taking into account precise frequency noise characteristic of the local oscillator, allowing to determine the optimal stationary operation regime of the clock (Macieszczak *et al.*, 2013) ideally in terms of the Allan variance taken as a figure of merit (Fraas, 2013). A deeper theoretical insight into this problem is still required to yield computable fundamental bounds.

The applicability of the tools presented has also been restricted to single phase parameter estimation. A more general approach may be taken, where multiple-phases (Humphreys *et al.*, 2013) or the phase as well as the decoherence strength itself are the quantities to be estimated (Crowley *et al.*, 2014; Knysh and Durkin, 2013). This poses an additional theoretical challenge as then the multi-parameter quantum estimation theory needs to be applied, while most of the tools discussed in this review are applicable only to single-parameter estimation. Developing non-trivial multi-parameter fundamental bounds for quantum metrology is therefore still an open field for research.

This research work supported by the FP7 IP project SIQS co-financed by the Polish Ministry of Science and Higher Education, Polish NCBiR under the ERA-NET CHIST-ERA project QUASAR, and Foundation for Polish Science TEAM project.

## REFERENCES

- Adesso, G. (2006), *Entanglement of Gaussian states*, Ph.D. thesis (Universita degli Studi di Salerno).
- Afek, I., O. Ambar, and Y. Silberberg (2010), *Science* **328** (5980), 879.
- Alipour, S. (2014), ArXiv e-prints arXiv:1403.8033 [quant-ph].
- Alipour, S., M. Mehboudi, and A. T. Rezakhani (2014), *Phys. Rev. Lett.* **112**, 120405.
- Anderson, B. P., and M. A. Kasevich (1998), *Science* **282** (5394), 1686.
- Andrè, A., A. S. Sørensen, and M. D. Lukin (2004), *Phys. Rev. Lett.* **92**, 230801.
- Anisimov, P. M., G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling (2010), *Phys. Rev. Lett.* **104**, 103602.
- Arrad, G., Y. Vinkler, D. Aharonov, and A. Retzker (2014), *Phys. Rev. Lett.* **112**, 150801.
- Aspachs, M., J. Calsamiglia, R. Muñoz Tapia, and E. Bagan (2009), *Phys. Rev. A* **79**, 033834.
- Bachor, H.-A., and T. Ralph (2004), *A guide to experiments in quantum optics* (Wiley-VCH).
- Banaszek, K., C. Radzewicz, K. Wodkiewicz, and J. S. Krasinski (1999), *Phys. Rev. A* **60**, 674.
- Banaszek, K., and K. Wódkiewicz (1998), *Phys. Rev. A* **58**, 4345.
- Barnett, S. M., and D. T. Pegg (1992), *Journal of Modern Optics* **39** (10), 2121.
- Bartlett, S. D., T. Rudolph, and R. W. Spekkens (2007), *Rev. Mod. Phys.* **79** (2), 555.
- Benatti, F., and D. Braun (2013), *Phys. Rev. A* **87**, 012340.
- Benatti, F., R. Floreanini, and U. Marzolino (2010), *Annals of Physics* **325** (4), 924.
- Bengtsson, I., and K. Życzkowski (2006), *Geometry of quantum states: an introduction to quantum entanglement* (Cambridge University Press).
- Berry, D. W., M. J. W. Hall, M. Zwierz, and H. M. Wiseman (2012), *Phys. Rev. A* **86**, 053813.
- Berry, D. W., B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman (2009), *Phys. Rev. A* **80** (5), 052114.
- Berry, D. W., and H. M. Wiseman (2000), *Phys. Rev. Lett.* **85** (24), 5098.
- Bollinger, J. J., W. M. Itano, D. J. Wineland, and D. J. Heinzen (1996), *Phys. Rev. A* **54** (6), R4649.
- Bondurant, R. S., and J. H. Shapiro (1984), *Phys. Rev. D* **30**, 2548.
- Boto, A. N., P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling (2000), *Phys. Rev. Lett.* **85**, 2733.
- Braunstein, S. L. (1992), *Phys. Rev. Lett.* **69**, 3598.
- Braunstein, S. L., and C. M. Caves (1994), *Phys. Rev. Lett.* **72** (22), 3439.
- Budker, D., and M. Romalis (2007), *Nature Phys.* **3**, 227.
- Bužek, V., R. Derka, and S. Massar (1999), *Phys. Rev. Lett.* **82**, 2207.
- Buzek, V., and P. L. Knight (1995) (Elsevier) pp. 1 – 158.
- Caves, C. M. (1981), *Phys. Rev. D* **23** (8), 1693.
- Chaves, R., J. B. Brask, M. Markiewicz, J. Kołodyński, and A. Acín (2013), *Phys. Rev. Lett.* **111**, 120401.
- Chekhova, M. (2011) (Elsevier) pp. 187 – 226.
- Chin, A. W., S. F. Huelga, and M. B. Plenio (2012), *Phys. Rev. Lett.* **109** (23), 233601, arXiv:1103.1219 [quant-ph].
- Chiribella, G., G. M. D’Ariano, and M. F. Sacchi (2005), *Phys. Rev. A* **72**, 042338.
- Choi, M.-D. (1975), *Linear Algebr. Appl.* **10**, 285.
- Chwedeńczuk, J., P. Hyllus, F. Piazza, and A. Smerzi (2012), *New Journal of Physics* **14** (9), 093001.
- Chwedeńczuk, J., F. Piazza, and A. Smerzi (2011), *New Journal of Physics* **13** (6), 065023.
- Cronin, A. D., J. Schmiedmayer, and D. E. Pritchard (2009), *Rev. Mod. Phys.* **81**, 1051.
- Crowley, P. J. D., A. Datta, M. Barbieri, and I. A. Walmsley (2014), *Phys. Rev. A* **89**, 023845.
- D’Ariano, G. M., C. Macchiavello, and M. G. Paris (1995), *Phys. Lett. A* **198** (4), 286.
- Datta, A., L. Zhang, N. Thomas-Peter, U. Dorner, B. J. Smith, and I. A. Walmsley (2011), *Phys. Rev. A* **83**, 063836.
- Demkowicz-Dobrzański, R. (2010), *Laser Physics* **20**, 1197.

- Demkowicz-Dobrzański, R., K. Banaszek, and R. Schnabel (2013), Phys. Rev. A **88**, 041802(R).
- Demkowicz-Dobrzański, R., U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley (2009), Phys. Rev. A **80**, 013825.
- Demkowicz-Dobrzański, R., J. Kołodyński, and M. Guţă (2012), Nat. Commun. **3**, 1063.
- Diddams, S. A., J. C. Bergquist, S. R. Jefferts, and C. W. Oates (2004), Science **306** (5700), 1318.
- Dorner, U. (2012), New J. Phys. **14** (4), 043011.
- Dorner, U., R. Demkowicz-Dobrzański, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley (2009), Phys. Rev. Lett. **102**, 040403.
- Dowling, J. P. (1998), Phys. Rev. A **57**, 4736.
- Dür, W., M. Skotiniotis, F. Fröwis, and B. Kraus (2014), Phys. Rev. Lett. **112**, 080801.
- Escher, B. M., L. Davidovich, N. Zagury, and R. L. de Matos Filho (2012), Phys. Rev. Lett. **109**, 190404.
- Escher, B. M., R. L. de Matos Filho, and L. Davidovich (2011), Nature Phys. **7**, 406.
- Fox, M. (2006), *Quantum Optics* (Ox).
- Fraas, M. (2013), ArXiv e-prints arXiv:1303.6083 [math-ph].
- Fröwis, F., M. Skotiniotis, B. Kraus, and W. Dür (2014), ArXiv e-prints arXiv:1402.6946.
- Fujiwara, A., and H. Imai (2008), J. Phys. A: Math. Theor. **41**, 255304.
- Genoni, M. G., S. Olivares, and M. G. A. Paris (2011), Phys. Rev. Lett. **106**, 153603.
- Gilbert, G., M. Hamrick, and Y. S. Weinstein (2008), J. Opt. Soc. Am B **25**, 1336.
- Giovannetti, V., S. Lloyd, and L. Maccone (2004), Science **306**, 1330.
- Giovannetti, V., S. Lloyd, and L. Maccone (2006), Phys. Rev. Lett. **96**, 010401.
- Giovannetti, V., S. Lloyd, and L. Maccone (2011), Nature Photon. **5**, 222.
- Giovannetti, V., and L. Maccone (2012), Phys. Rev. Lett. **108**, 210404.
- Glauber, R. J. (1963), Phys. Rev. **131**, 2766.
- Gross, C., T. Zibold, E. Nicklas, J. Esteve, and M. K. Oberthaler (2010), Nature **464**, 1165.
- Hall, M. J. W., and H. M. Wiseman (2012), New J. Phys. **14**, 033040.
- Hariharan, P. (2003), *Optical interferometry* (Elsevier, Amsterdam).
- Helstrom, C. W. (1976), *Quantum detection and estimation theory* (Academic press).
- Higgins, B. L., D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde (2007), Nature **450**, 393.
- Hofheinz, M., E. M. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O’Connell, H. Wang, J. M. Martinis, and A. N. Cleland (2008), Nature **454**, 310.
- Hofmann, H. F. (2009), Phys. Rev. A **79** (3), 033822.
- Holevo, A. S. (1982), *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam).
- Holland, M. J., and K. Burnett (1993), Phys. Rev. Lett. **71** (9), 1355.
- Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki (2009), Rev. Mod. Phys. **81**, 865.
- Horrom, T., R. Singh, J. P. Dowling, and E. E. Mikhailov (2012), Phys. Rev. A **86**, 023803.
- Hradil, Z., R. Myška, J. Peřina, M. Zawisky, Y. Hasegawa, and H. Rauch (1996), Phys. Rev. Lett. **76**, 4295.
- Huelga, S. F., C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac (1997), Phys. Rev. Lett. **79** (20), 3865.
- Humphreys, P. C., M. Barbieri, A. Datta, and I. A. Walmsley (2013), Phys. Rev. Lett. **111**, 070403.
- Huuer, S. D., C. F. Wildfeuer, and J. P. Dowling (2008), Phys. Rev. A **78**, 063828.
- Hyllus, P., W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi (2012), Phys. Rev. A **85**, 022321.
- Jamiołkowski, A. (1972), Rep. Math. Phys **3**, 275.
- Jarzyna, M., and R. Demkowicz-Dobrzański (2012), Phys. Rev. A **85**, 011801(R).
- Jarzyna, M., and R. Demkowicz-Dobrzański (2013), Phys. Rev. Lett. **110**, 240405.
- Jarzyna, M., and R. Demkowicz-Dobrzański (2014), “in preparation”.
- Jeske, J., J. H. Cole, and S. Huelga (2013), ArXiv e-prints arXiv:1307.6301 [quant-ph].
- Joo, J., W. J. Munro, and T. P. Spiller (2011), Phys. Rev. Lett. **107**, 083601.
- Kacprowicz, M., R. Demkowicz-Dobrzański, W. Wasilewski, and K. Banaszek (2010), Nature Photon. **4**, 357.
- Kaftal, T., and R. Demkowicz-Dobrzański (2014), ArXiv e-prints arXiv:1405.5897 [quant-ph].
- Kay, S. M. (1993), *Fundamentals of Statistical Signal Processing: Estimation Theory*, edited by S. E. Alan V. Oppenheim (Prentice Hall).
- Kessler, E. M., I. Lovchinsky, A. O. Sushkov, and M. D. Lukin (2013), ArXiv e-prints arXiv:1310.3260 [quant-ph].
- Killoran, N., M. Cramer, and M. B. Plenio (2014), Phys. Rev. Lett. **112**, 150501.
- Kitagawa, M., and M. Ueda (1993), Phys. Rev. A **47**, 5138.
- Knysh, S., V. N. Smelyanskiy, and G. A. Durkin (2011), Phys. Rev. A **83**, 021804.
- Knysh, S. I., E. H. Chen, and G. A. Durkin (2014), ArXiv e-prints arXiv:1402.0495 [quant-ph].
- Knysh, S. I., and G. A. Durkin (2013), ArXiv e-prints arXiv:1307.0470 [quant-ph].
- Kok, P., W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn (2007), Rev. Mod. Phys. **79**, 135.
- Kołodyński, J., and R. Demkowicz-Dobrzański (2010), Phys. Rev. A **82**, 053804.
- Kołodyński, J., and R. Demkowicz-Dobrzański (2013), New J. Phys. **15**, 073043.
- Krischek, R., C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi (2011), Phys. Rev. Lett. **107**, 080504.
- Lee, H., P. Kok, and J. P. Dowling (2002), Journal of Modern Optics **49** (14-15), 2325.
- Lehmann, E. L., and G. Casella (1998), *Theory of point estimation*, Vol. 31 (Springer).
- Leibfried, D., M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland (2004), Science **304**, 1476.
- Leonhardt, U. (1997), *Measuring the quantum state of light*, Vol. 22 (Cambridge University Press).
- LIGO Collaboration, (2011), Nature Phys. **7**, 962.
- LIGO Collaboration, (2013), Nature Photon. **7** (8), 613.
- Loudon, R. (2000), *The quantum theory of light* (Oxford University Press).
- Luis, A., and J. Perina (1996), Phys. Rev. A **54**, 4564.
- Lynch, R. (1995), Phys. Rep. **256** (6), 367.
- Ma, J., X. Wang, C. Sun, and F. Nori (2011), Phys. Rep. **509** (2), 89.

- Macieszczak, K. (2014), ArXiv e-prints arXiv:1403.0955 [quant-ph].
- Macieszczak, K., R. Demkowicz-Dobrzański, and M. Fraas (2013), ArXiv e-prints arXiv:1311.5576 [quant-ph].
- Mandel, L., and E. Wolf (1995), *Optical coherence and quantum optics* (Camb).
- Matsumoto, K. (2010), ArXiv e-prints , 1006.0300v1.
- Matsuzaki, Y., S. C. Benjamin, and J. Fitzsimons (2011), Phys. Rev. A **84**, 012103.
- Mitchell, M. W., J. S. Lundeen, and A. M. Steinberg (2004), Nature **429**, 161.
- Mølmer, K. (1997), Phys. Rev. A **55**, 3195.
- Nagaoka, H. (2005), “On fisher information of quantum statistical models,” Chap. 9 (World) pp. 113–125.
- Nagata, T., R. Okamoto, J. L. O’Brien, K. Sasaki, and S. Takeuchi (2007), Science **316**, 726.
- Nair, R. (2012), ArXiv e-prints arXiv:1204.3761 [quant-ph].
- Nielsen, M. A., and I. L. Chuang (2000), *Quantum Computing and Quantum Information* (Cambridge University Press).
- Noh, J. W., A. Fougères, and L. Mandel (1992), Phys. Rev. A **45**, 424.
- Okamoto, R., H. F. Hofmann, T. Nagata, J. L. O’Brien, K. Sasaki, and S. Takeuchi (2008), New Journal of Physics **10** (7), 073033.
- Olivares, S., and M. G. A. Paris (2007), Opt. Spectrosc. **103**, 231.
- Ono, T., and H. F. Hofmann (2010), Phys. Rev. A **81**, 033819.
- Pan, J.-W., Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski (2012), Rev. Mod. Phys. **84**, 777.
- Pezzé, L., and A. Smerzi (2008), Phys. Rev. Lett. **100** (7), 073601.
- Pezzé, L., and A. Smerzi (2009), Phys. Rev. Lett. **102**, 100401.
- Pinel, O., J. Fader, D. Braun, P. Jian, N. Treps, and C. Fabre (2012), Phys. Rev. A **85**, 010101(R).
- Pinel, O., P. Jian, C. Fabre, N. Treps, and D. Braun (2013), Phys. Rev. A **88**, 040102(R).
- Pitkin, M., S. Reid, S. Rowan, and J. Hough (2011), Living Reviews in Relativity **14** (5), 10.12942/lrr-2011-5.
- Reid, M. D., and P. D. Drummond (1988), Phys. Rev. Lett. **60**, 2731.
- Resch, K. J., K. L. Pregnell, R. Prevedel, A. Gilchrist, G. J. Pryde, J. L. O’Brien, and A. G. White (2007), Phys. Rev. Lett. **98** (22), 223601.
- Rivas, A., and A. Luis (2012), New Journal of Physics **14** (9), 093052.
- Rubin, M. A., and S. Kaushik (2007), Phys. Rev. A **75**, 053805.
- Sanders, B. C., and G. J. Milburn (1995), Phys. Rev. Lett. **75**, 2944.
- Sarovar, M., and G. J. Milburn (2006), J. Phys. A: Math. Gen. **39**, 8487.
- Sayrin, C., I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, J.-M. Raimond, and S. Haroche (2011), Nature **477**, 73.
- Schleich, W. P. (2001), *Quantum optics in phase space* (Wiley-VCH).
- Schwinger, J. (1965), “Quantum theory of angular momentum,” (Academic Press, New York).
- Seshadreesan, K. P., P. M. Anisimov, H. Lee, and J. P. Dowling (2011), New Journal of Physics **13**, 083026.
- Sewell, R. J., M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell (2012), Phys. Rev. Lett. **109**, 253605.
- Shaji, A., and C. M. Caves (2007), Phys. Rev. A **76**, 032111.
- Shankar, R. (1994), *Principles of Quantum Mechanics* (Plenum).
- Shin, Y., M. Saba, T. A. Pasquini, W. Ketterle, D. E. Pritchard, and A. E. Leanhardt (2004), Phys. Rev. Lett. **92**, 050405.
- Spagnolo, N., C. Vitelli, V. G. Lucivero, V. Giovannetti, L. Maccone, and F. Sciarrino (2012), Phys. Rev. Lett. **108** (23), 233602.
- Summy, G. S., and D. T. Pegg (1990), Opt. Commun. **77** (1), 75.
- Torres, J. P., K. Banaszek, and I. Walmsley (2011) (Elsevier) pp. 227 – 331.
- Tóth, G. (2012), Phys. Rev. A **85**, 022322.
- Tóth, G., and D. Petz (2013), Phys. Rev. A **87**, 032324.
- Tsang, M. (2012), Phys. Rev. Lett. **108**, 230401.
- van der Vaart, A. W. (2000), *Asymptotic Statistics* (Cambridge).
- Vitelli, C., N. Spagnolo, L. Toffoli, F. Sciarrino, and F. De Martini (2010), Phys. Rev. Lett. **105** (11), 113602.
- Walls, D. F., and G. J. Milburn (1995), *Quantum Optics* (Springer-Verlag, Berlin).
- Wasilewski, W., K. Jensen, H. Krauter, J. J. Renema, M. V. Balabas, and E. S. Polzik (2010), Phys. Rev. Lett. **104**, 133601.
- Wigner, E. P. (1932), Phys. Rev. **40**, 749.
- Wineland, D. J., J. J. Bollinger, W. M. Itano, and D. J. Heinzen (1994), Phys. Rev. A **50**, 67.
- Wineland, D. J., J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen (1992), Phys. Rev. A **46**, R6797.
- Xiang, G.-Y., B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde (2010), Nature Photon. **5**, 43.
- Xiao, M., L.-A. Wu, and H. J. Kimble (1987), Phys. Rev. Lett. **59** (3), 278.
- Yu, S. (2013), ArXiv e-prints arXiv:1302.5311 [quant-ph].
- Yurke, B., S. L. McCall, and J. R. Klauder (1986), Phys. Rev. A **33** (6), 4033.
- Zhang, Y. R., G. R. Jin, J. P. Cao, W. M. Liu, and H. Fan (2013), J. Phys. A: Math. Theor. **46**, 035302.
- Zwierz, M., C. A. Pérez-Delgado, and P. Kok (2010), Phys. Rev. Lett. **105**, 180402.